

# WELL-POSEDNESS OF EVOLUTION EQUATIONS WITH TIME-DEPENDENT NONLINEAR MOBILITY: A MODIFIED MINIMIZING MOVEMENT SCHEME

JONATHAN ZINSL

**ABSTRACT.** We prove the existence of nonnegative weak solutions to a class of second and fourth order nonautonomous nonlinear evolution equations with an explicitly time-dependent mobility function posed on the whole space  $\mathbb{R}^d$ , for arbitrary  $d \geq 1$ . Exploiting a very formal gradient flow structure, the cornerstone of our proof is a modified version of the classical minimizing movement scheme for gradient flows. The mobility function is required to satisfy—at each time point separately—the conditions by which one can define a modified Wasserstein distance on the space of probability densities with finite second moment. The explicit dependency on the time variable is assumed to be at least of Lipschitz regularity. We also sketch possible extensions of our result to the case of bounded spatial domains and more general mobility functions.

## 1. INTRODUCTION

In this article, the following nonautonomous partial differential equation is considered:

$$\partial_t u(t, x) = \operatorname{div} \left( m(t, u(t, x)) \nabla \frac{\delta \mathcal{E}}{\delta u}(u(t, x)) \right), \quad t > 0, x \in \mathbb{R}^d, \quad (1.1)$$

together with the initial condition  $u(0, \cdot) = u_0$  on  $\mathbb{R}^d$  for  $d \geq 1$ . In addition, we seek for a *nonnegative* solution  $u(t, x) \geq 0$  for all  $t \geq 0$  and almost all  $x \in \mathbb{R}^d$ .

Above,  $\frac{\delta \mathcal{E}}{\delta u}$  denotes the first variation of the *free energy* functional  $\mathcal{E}$  in  $L^2$ . Our assumptions on  $\mathcal{E}$  are specified below in the conditions (E1) and (E2), respectively. There, we distinguish two main classes of energy functionals. First, in the case (E1),  $\mathcal{E}$  is assumed to be of the form

$$\mathcal{E}(u) = \int_{\mathbb{R}^d} [f(u(x)) + \phi(x)u(x)] \, dx, \quad (1.2)$$

where the internal energy density  $f$  is uniformly convex in  $u$ , and the *confinement potential*  $\phi$  is bounded from below and grows at most quadratically as  $|x| \rightarrow \infty$ . Equation (1.1) then reads as a nonlinear second order drift-diffusion equation,

$$\partial_t u = \operatorname{div} [m(t, u)(f''(u)\nabla u + \nabla \phi)]. \quad (1.3)$$

One may also allow for *gradient-dependent* energy density:

$$\mathcal{E}(u) = \int_{\mathbb{R}^d} [f(\nabla u(x), u(x)) + \phi(x)u(x)] \, dx, \quad (1.4)$$

where again  $f$  is assumed to satisfy a certain convexity condition. Equation (1.1) then is the fourth-order equation

$$\partial_t u = \operatorname{div} [m(t, u)\nabla(-\operatorname{div} \nabla_p f(\nabla u, u) + \partial_z f(\nabla u, u) + \phi)]. \quad (1.5)$$

---

*Date:* April 27, 2016.

2010 *Mathematics Subject Classification.* Primary: 35K30; Secondary: 35A15, 35D30.

*Key words and phrases.* Non-autonomous equation; gradient flow; nonlinear mobility; modified Wasserstein distance; minimizing movement scheme.

This research has been supported by the DFG Collaborative Research Center TRR 109 “Discretization in Geometry and Dynamics”.

We do not seek for full generality of energy functionals  $\mathcal{E}$  here, since our main interest is focussed on the *time-dependent mobility function*  $m : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  which turns (1.1) into a *non-autonomous* nonlinear evolution equation.

When  $m$  does not explicitly depend on  $t$ , it is known that (1.1) possesses a variational structure [8, 16, 17], if  $m$  is nonnegative and concave on the interior of an interval  $[0, S]$ , the so-called *value space* [29], with  $S \in \mathbb{R}_{\geq 0} \cup \{+\infty\}$ . In this work, we require that at each fixed time  $t \geq 0$ ,  $m(t, \cdot)$  is an admissible mobility in the sense of [8, 16] (see condition (M1) below), where the corresponding value spaces  $[0, S(t)]$  are assumed to be expanding over time, i.e.,  $S$  is nondecreasing. For the dependency on the time variable, certain regularity—at least Lipschitz—conditions are needed. The detailed assumptions on  $m$  are presented in (M1)–(M4) below.

**Example 1.1** (Paradigmatic examples). The following mobility functions admit the conditions (M1)–(M4) below:

(a) *Finite value spaces*: For all  $t \geq 0$ , put

$$m(t, z) = z(S(t) - z) \quad \text{for each } z \in [0, S(t)],$$

for some sufficiently regular nondecreasing function  $S$ . Choosing the quadratic free energy

$$\mathcal{E}_Q(u) = \frac{1}{2} \int_{\mathbb{R}^d} |\nabla u(x)|^2 dx + \frac{1}{2} \int_{\mathbb{R}^d} u(x)^2 dx, \quad (1.6)$$

which is admissible in (E2), the fourth-order equation (1.5) reads as a variant of the *Cahn-Hilliard* equation,

$$\partial_t u = -\operatorname{div}(u(S(t) - u)\nabla \Delta u) + \operatorname{div}(u(S(t) - u)\Delta u).$$

(b) *Infinite value space*, i.e.,  $S \equiv +\infty$ : Let  $\varepsilon > 0$  and let  $\alpha : \mathbb{R}_{\geq 0} \rightarrow (0, 1]$  be sufficiently regular. Set

$$m(t, z) = (z + \varepsilon)^{\alpha(t)} - \varepsilon^{\alpha(t)} \quad \text{for all } z \geq 0.$$

With the quadratic free energy  $\mathcal{E}_Q$  from (1.6), (1.5) is a perturbed version of the *thin-film* equation,

$$\partial_t u = -\operatorname{div}(u((u + \varepsilon)^{\alpha(t)} - \varepsilon^{\alpha(t)})\nabla \Delta u) + \operatorname{div}(u((u + \varepsilon)^{\alpha(t)} - \varepsilon^{\alpha(t)})\Delta u).$$

Notice that in (b), the case  $\varepsilon = 0$  is (whenever  $\alpha$  is not identically equal to 1) not allowed since condition (M3) requires Lipschitz-continuity of  $m(t, \cdot)$  also in the state variable  $z$ . For a possible generalization where (M3) can be dispensed of, see Section 5.

We use the (pointwise in  $t$ ) formal gradient structure of (1.1) from [8, 16] on the space

$$\mathbf{X}(t) := \left\{ u \in L^1(\mathbb{R}^d) : 0 \leq u(x) \leq S(t) \text{ a.e. on } \mathbb{R}^d, \|u\|_{L^1} = 1 \text{ and } \int_{\mathbb{R}^d} |x|^2 u(x) dx < \infty \right\}, \quad (1.7)$$

with respect to the metric  $\mathbf{W}_{m(t, \cdot)}$  induced by  $m(t, \cdot)$ :

$$\begin{aligned} & \mathbf{W}_{m(t, \cdot)}(u_0, u_1) \\ &= \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} \frac{|w_s|^2}{m(t, u_s)} dx ds : (u_s, w_s)_{s \in [0, 1]} \in \mathcal{C}, u_s|_{s=0} = u_0, u_s|_{s=1} = u_1 \right\}^{1/2}, \end{aligned} \quad (1.8)$$

where  $\mathcal{C}$  is a suitable subclass of solutions to the *continuity equation*  $\partial_s u_s = -\operatorname{div} w_s$  in the sense of distributions on  $[0, 1] \times \mathbb{R}^d$  (see [8, 16] for more details). By convention, we set  $\mathbf{W}_{m(t, \cdot)}(u_0, u_1) = +\infty$  if  $u_0$  and  $u_1$  are not elements of  $\mathbf{X}(t)$  both. If  $m(t, \cdot)$  is a linear function, we recover (a scalar multiple of) the classical  $L^2$ -Wasserstein distance  $\mathbf{W}_2$  on the space of probability measures [4] (see Section 2 below).

Specifically, we prove the existence of nonnegative weak solutions to (1.1) using a modified version of the classical *minimizing movement scheme* for gradient flows (which has been employed for various evolution equations with metric gradient flow structure, e.g. [11, 21, 1, 10, 19, 12, 17, 27, 29, 5], also with spatially varying coefficients [22, 15, 18]):

**Definition 1.2** (Modified minimizing movement scheme). Define the space

$$\mathbf{X} := \{u \in L^1(\mathbb{R}^d) : 0 \leq u(x) \text{ a.e. on } \mathbb{R}^d, \|u\|_{L^1} = 1 \text{ and } \mathbf{m}_2(u) < \infty\}, \quad (1.9)$$

where  $\mathbf{m}_2(u) = \int_{\mathbb{R}^d} |x|^2 u(x) dx$  denotes the second moment of  $u$ .

Let now  $\tau > 0$  and  $u_0 \in \mathbf{X}$  be given. A sequence  $(u_\tau^n)_{n \in \mathbb{N}}$  is obtained via the *modified minimizing movement scheme* if  $u_\tau^0 = u_0$  and for all  $n \in \mathbb{N}$ ,

$$u_\tau^n \in \operatorname{argmin}_{u \in \mathbf{X}} \left[ \frac{1}{2\tau} \mathbf{W}_{m(n\tau, \cdot)}(u_\tau^{n-1}, u)^2 + \mathcal{E}(u) \right]. \quad (1.10)$$

The *discrete solution*  $u_\tau : \mathbb{R}_{\geq 0} \times \mathbf{X} \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$  is defined by piecewise constant interpolation along the sequence  $(u_\tau^n)_{n \in \mathbb{N}}$ , that is

$$u_\tau(t) := u_\tau^n, \quad \text{for } n := \left\lceil \frac{t}{\tau} \right\rceil \quad \text{for all } t \geq 0.$$

We now present our main assumptions and results.

**1.1. Assumptions and main results.** For the mobility function, we require the following.

**Definition 1.3.** (Admissible mobility functions)

(M1) There exists a nondecreasing function  $S : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{>0} \cup \{+\infty\}$  such that for fixed  $t \geq 0$ , the map  $m(t, \cdot) \in C^2((0, S(t)))$  is an admissible mobility for the definition of a distance  $W_{m(t, \cdot)}$ , viz.

$$\begin{aligned} m(t, z) &> 0 \quad \text{for all } z \in (0, S(t)), \quad \text{and} \quad m(t, z) = 0 \quad \text{for all } z \in \mathbb{R}_{\geq 0} \setminus (0, S(t)), \\ \partial_z^2 m(t, z) &\leq 0 \quad \text{for all } z \in (0, S(t)). \end{aligned}$$

(M2) For fixed  $z \geq 0$ , the map  $m(\cdot, z)$  is locally Lipschitz-continuous on  $\mathbb{R}_{\geq 0}$ .

(M3) There exist  $M_1, M_2 \in C(\mathbb{R}_{\geq 0}; \mathbb{R}_{>0})$  such that for all  $t \geq 0$ :

$$\sup_{z \in (0, S(t))} |\partial_z m(t, z)| \leq M_1(t), \quad \sup_{z \in (0, S(t))} (-m(t, z) \partial_z^2 m(t, z)) \leq M_2(t).$$

(M4) There exists a map  $h : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  such that for fixed  $t \geq 0$ ,  $m(t, \cdot)$  is *induced* by  $h(t, \cdot)$ , i.e.,

$$m(t, z) \partial_z^2 h(t, z) = 1 \quad \text{for all } z \in (0, S(t)).$$

The map  $h$  is assumed to admit the following bound:

For fixed  $t \geq 0$  and all  $z \in [0, S(t))$ , one has

$$|h(t, z)| \leq H_1(t) z^{\alpha_1(t)} + H_2(t) z^{\alpha_2(t)+1},$$

where  $H_1(t), H_2(t) \geq 0$ ,  $\alpha_1(t) \in \left(\frac{d}{d+2}, 1\right)$  and  $\alpha_2(t) \in [0, 1]$  all depend continuously on  $t$ .

$h$  is locally Lipschitz w.r.t.  $t$  in the following sense:

For all  $T > 0$ , all  $0 \leq t \leq t' \leq T$  and all  $z \in [0, S(t)]$ :

$$|h(t', z) - h(t, z)| \leq (L_1(T) z^{\beta_1(T)} + L_2(T) z^{\beta_2(T)+1})(t' - t),$$

where  $L_1(T), L_2(T) \geq 0$ ,  $\beta_1(T) \in \left(\frac{d}{d+2}, 1\right)$  and  $\beta_2(T) \in [0, 1]$  all depend continuously on  $T$ .

The conditions (M1) and (M3) are concerned with the behaviour of  $m$  for fixed time variable  $t$ : whereas (M1) is needed to be able to define a metric  $\mathbf{W}_{m(t, \cdot)}$  on  $\mathbf{X}(t)$ , assumption (M3)—which is a time-dependent version of the so-called *Lipschitz-semiconcavity (LSC)* condition from [17] (see also [28])—yields  $\lambda$ -convexity along geodesics of the viscous regularization of the potential energy  $u \mapsto \int_{\mathbb{R}^d} u \eta dx$  for some  $\lambda \leq 0$ . This will be a cornerstone in our derivation of a time-discrete approximate weak formulation satisfied by the discrete solution  $u_\tau$ . In contrast, (M2) and (M4)

are concerned with the behaviour of  $m$  w.r.t. time  $t$ . There exists, at each  $t \geq 0$  separately, one distinguished 0-geodesically convex functional  $\mathcal{H}_t$ ,

$$\mathcal{H}_t(u) := \int_{\mathbb{R}^d} h(t, u(x)) \, dx,$$

the so-called *heat entropy* [29] since it induces the heat flow as its  $\mathbf{W}_{m(t, \cdot)}$ -gradient flow [8, 16, 17, 29]. Assumption (M4) is used to control the behaviour of  $\mathcal{H}_t$  with respect to time  $t$ , guaranteeing sufficient spatial regularity of the (time-)discrete solution  $u_\tau$ .

Our energy functionals are assumed to satisfy one of the following two conditions.

**Definition 1.4** (Admissible energy functionals).

- (E1) *Second order equations* (1.3): Let  $\mathcal{E}$  be of the form (1.2), where  $f \in C^2(\mathbb{R}_{\geq 0})$  with  $f(0) = 0$ ,  $f'(0) = 0$  and  $f''(z) \in [\gamma_0, \gamma_1]$  for  $0 < \gamma_0 \leq \gamma_1 < \infty$  and all  $z \geq 0$ , and  $\phi \in C^2(\mathbb{R}^d)$  is bounded from below and grows at most quadratically, with  $\Delta\phi \in L^\infty(\mathbb{R}^d)$ .
- (E2) *Fourth order equations* (1.5): Let  $\mathcal{E}$  be of the form (1.4), with  $\phi$  as in (E1) and  $f \in C^2(\mathbb{R}^d \times \mathbb{R}_{\geq 0})$ ,  $f(0, 0) = 0$ ,  $\nabla_{(p, z)} f(0, 0) = 0$  and  $\gamma_0 \mathbb{1}_{d+1} \leq \nabla_{(p, z)}^2 f(p, z) \leq \gamma_1 \mathbb{1}_{d+1}$  for all  $p \in \mathbb{R}^d$  and all  $z \geq 0$ , with  $0 < \gamma_0 \leq \gamma_1 < \infty$ .

We obtain the following on the convergence behaviour of  $u_\tau$  as  $\tau \searrow 0$ .

**Theorem 1.5** (Existence: second order case (E1)). *Assume that the mobility satisfies (M1)–(M4) and the energy functional  $\mathcal{E}$  is of the form (E1) and let an initial datum  $u_0 \in \mathbf{X} \cap L^2(\mathbb{R}^d)$  with  $u_0(x) \in [0, S(0)]$  for almost every  $x \in \mathbb{R}^d$  be given. Then, for each  $\tau > 0$ , the map  $u_\tau$  obtained via (1.10) is well-defined. Furthermore, for each vanishing sequence  $\tau_k \searrow 0$  ( $k \rightarrow \infty$ ), there exists a (nonrelabelled) subsequence and a limit map  $u : \mathbb{R}_{\geq 0} \rightarrow \mathbf{X}$  such that the following is true for each fixed  $T > 0$ :*

- (a)  $u \in L^\infty([0, T]; L^2(\mathbb{R}^d)) \cap L^2([0, T]; H^1(\mathbb{R}^d)) \cap C^{1/2}([0, T]; \mathbf{W}_2)$ , and for almost every  $t \geq 0$ , one has  $u(t, x) \leq S(t)$  for almost all  $x \in \mathbb{R}^d$ ;
- (b)  $u_{\tau_k}(t) \rightarrow u(t)$  weakly\* in  $\mathcal{P}$ , at each fixed  $t > 0$ ;
- (c)  $u_{\tau_k}$  converges to  $u$  strongly in  $L^2([0, T]; L^2(\Omega))$  for each bounded domain  $\Omega \subset \mathbb{R}^d$  and weakly in  $L^2([0, T]; H^1(\mathbb{R}^d))$ ;
- (d)  $u$  is a solution to (1.3) in the sense of distributions and  $u(0, \cdot) = u_0$  a.e. on  $\mathbb{R}^d$ .

A similar statement also holds in the fourth order case.

**Theorem 1.6** (Existence: fourth order case (E2)). *Assume that the mobility satisfies (M1)–(M4) and the energy functional  $\mathcal{E}$  is of the form (E2) and let an initial datum  $u_0 \in \mathbf{X} \cap H^1(\mathbb{R}^d)$  with  $u_0(x) \in [0, S(0)]$  for almost every  $x \in \mathbb{R}^d$  be given. Then, for each  $\tau > 0$ , the map  $u_\tau$  obtained via (1.10) is well-defined. Furthermore, for each vanishing sequence  $\tau_k \searrow 0$  ( $k \rightarrow \infty$ ), there exists a (nonrelabelled) subsequence and a limit map  $u : \mathbb{R}_{\geq 0} \rightarrow \mathbf{X}$  such that the following is true for each fixed  $T > 0$ :*

- (a)  $u \in L^\infty([0, T]; H^1(\mathbb{R}^d)) \cap L^2([0, T]; H^2(\mathbb{R}^d)) \cap C^{1/2}([0, T]; \mathbf{W}_2)$ , and for almost every  $t \geq 0$ , one has  $u(t, x) \leq S(t)$  for almost all  $x \in \mathbb{R}^d$ ;
- (b)  $u_{\tau_k}(t) \rightarrow u(t)$  weakly\* in  $\mathcal{P}$ , at each fixed  $t > 0$ ;
- (c)  $u_{\tau_k}$  converges to  $u$  strongly in  $L^2([0, T]; H^1(\Omega))$  for each bounded domain  $\Omega \subset \mathbb{R}^d$  and weakly in  $L^2([0, T]; H^2(\mathbb{R}^d))$ ;
- (d)  $u$  is a solution to (1.5) in the sense of distributions and  $u(0, \cdot) = u_0$  a.e. on  $\mathbb{R}^d$ .

Note that in both cases, we do neither obtain uniqueness of solutions nor a monotonicity property of the energy functional in the limit  $\tau \searrow 0$ , since our notion of solution is very weak and the minimization problem in (1.10) lacks convexity. Notice furthermore that our results in principle also hold on a bounded and convex spatial domain. We refer to Section 5 for a sketch of possible extensions in that direction.

**1.2. Strategy of proof and relation to the literature.** Our strategy of proof can be summarized as follows. Given the conditions (E1)/(E2) and (M1)–(M4), the map  $u_\tau$  obtained via the modified minimizing movement scheme (1.10) is well-defined. It obeys a certain regularity property which is crucial for the passage to the limit as  $\tau \searrow 0$  afterwards, since, due to the nonlinearity of the problem, weak convergence alone is not sufficient: there, we apply the so-called *flow interchange* technique introduced in [19]. Similarly as in [17, 29], the necessary auxiliary flow is given as the  $\mathbf{W}_{m(n\tau, \cdot)}$ -gradient flow of the *(time-dependent) heat entropy* functional

$$\mathcal{H}_{n\tau}(u) = \int_{\mathbb{R}^d} h(n\tau, u) \, dx,$$

where  $h$  is the function from (M4). This auxiliary flow is—by construction—the heat flow. Nonautonomous evolution equations of gradient flow type have already been studied in [9, 23] from the opposite point of view: there, time-dependent energy functionals on time-independent metric spaces were considered and a different modification of the minimizing movement scheme was investigated. In [23], where general non-convex problems have been studied, a certain Lipschitz condition for the free energy functional comes into play which resembles our additional conditions on  $h$  from (M4). In [9, 23], re-proving the classical properties of the minimizing movement scheme is more involved than for our scheme from Definition 1.2—in contrast, deriving higher regularity estimates brings additional difficulties in our case, since now the heat entropy is a time-dependent functional. Non-autonomous equations of Wasserstein gradient flow form with linear mobility have also been investigated in [22] using a time-averaged form of the classical minimizing movement scheme. Notice that our equations do not in general allow for a time-dependent scaling which transforms the problem into autonomous form (for studies in this direction, see e.g. [3]).

Nonautonomous linear and semilinear equations have, in contrast, been investigated more exhaustively using semigroup theory (see e.g. [2] and references therein). More recently, properties of time-dependent Riemannian manifolds have been studied [25], and corresponding evolution problems have been investigated [13].

The partial Riemannian structure on the space  $\mathbf{X}(t)$  induced by the distance  $\mathbf{W}_{m(t, \cdot)}$  was introduced in [8, 16] and later generalized to the vector-valued framework in [29]; the structure of geodesics on that space was first investigated in [6]. In [17], the formal gradient flow with respect to that distance was used to prove existence of solutions for a certain class of Cahn-Hilliard type fourth-order equations. There, also mobility functions which are not Lipschitz continuous w.r.t.  $z$  could be used using an approximation technique. In this work (see also [28]), we will employ this strategy to extend our results from Section 1.1 to wider classes of mobility functions as those satisfying (M1)–(M4), see Section 5.

**1.3. Plan of the paper.** The paper is organized as follows. In Section 2, we summarize some preliminary facts on gradient flows with respect to the  $L^2$ -Wasserstein distance and its modifications. Afterwards, the variational scheme from Definition 1.2 is studied in Section 3. Section 4 then is concerned with the derivation of the approximate discrete weak formulation and the passage to the continuous-time limit  $\tau \searrow 0$ . We sketch several possible generalizations in Section 5.

## 2. PRELIMINARIES

Derivatives with respect to state variables  $z$  or  $p$  will always be indicated as  $\partial_z$  or  $\nabla_p$ , respectively. Spatial derivatives  $\nabla$ ,  $\operatorname{div}$  or  $\Delta$  are denoted without further indication. We now briefly summarize some results from the theory of gradient flows in spaces of measures. Our ambient space  $\mathbf{X}$  from (1.9) is—by the slight abuse of notation where we identify probability densities  $u$  with their corresponding probability measures  $u \cdot \mathcal{L}^d$ —a subspace of the space of probability measures  $\mathcal{P}$  on  $\mathbb{R}^d$ . A sequence  $(\mu_k)_{k \in \mathbb{N}}$  in  $\mathcal{P}$  is said to converge *weakly\** to some  $\mu \in \mathcal{P}$  if for all continuous and bounded maps

$g : \mathbb{R}^d \rightarrow \mathbb{R}$ , one has

$$\lim_{k \rightarrow \infty} \int_{\mathbb{R}^d} g \, d\mu_k = \int_{\mathbb{R}^d} g \, d\mu.$$

The space  $\mathcal{P}$  can be endowed with the so-called  $L^2$ -Wasserstein (pseudo-)distance: For each  $\mu_0, \mu_1 \in \mathcal{P}$ ,

$$\mathbf{W}_2(\mu_0, \mu_1) = \inf \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y| \, d\gamma(x, y) : \gamma \in \Gamma(\mu_0, \mu_1) \right\}^{1/2},$$

where  $\Gamma(\mu_0, \mu_1)$  denotes the set of transport plans from  $\mu_0$  to  $\mu_1$  (for more details on optimal transport, see [26]). A dynamical characterization of  $\mathbf{W}_2$ —which will be made use of here—was found by Benamou and Brenier [4]:

$$\mathbf{W}_2(u_0, u_1) = \inf \left\{ \int_0^1 \int_{\mathbb{R}^d} \frac{|w_s|^2}{u_s} \, dx \, ds : (u_s, w_s)_{s \in [0,1]} \in \mathcal{C}, u_s|_{s=0} = u_0, u_s|_{s=1} = u_1 \right\}^{1/2},$$

where  $\mathcal{C}$  is a suitable subclass of solutions to the *continuity equation*  $\partial_s u_s = -\operatorname{div} w_s$  in the sense of distributions on  $[0, 1] \times \mathbb{R}^d$ . This characterization has been the starting point for the definition of the distances  $\mathbf{W}_{m(t, \cdot)}$  in (1.8) for *nonlinear* mobility functions [8, 16].

In order to derive the necessary *a priori* estimates on the discrete solution  $u_\tau$  obtained via our variational scheme from Definition 1.2, we use the so-called *flow interchange* lemma from [19]. In advance of its precise statement, we introduce the following notion of gradient flow (compare [1, 7]).

**Definition 2.1** ( $\kappa$ -flows). Let  $\mathcal{A} : \mathcal{P} \rightarrow \mathbb{R} \cup \{+\infty\}$  be a proper and lower semicontinuous functional on the (pseudo-)metric space  $(\mathcal{P}, \mathbf{d})$  and let  $\kappa \in \mathbb{R}$ . A continuous semigroup  $\mathbf{S}^{\mathcal{A}}$  on  $(\mathcal{P}, \mathbf{d})$  satisfying the *evolution variational estimate*

$$\frac{1}{2} \frac{d^+}{ds} \mathbf{d}^2(\mathbf{S}_s^{\mathcal{A}}(w), \tilde{w}) + \frac{\kappa}{2} \mathbf{d}^2(\mathbf{S}_s^{\mathcal{A}}(w), \tilde{w}) + \mathcal{A}(\mathbf{S}_s^{\mathcal{A}}(w)) \leq \mathcal{A}(\tilde{w})$$

for arbitrary  $w, \tilde{w} \in \operatorname{Dom}(\mathcal{A})$  and for all  $s \geq 0$ , as well as the monotonicity condition

$$\mathcal{A}(\mathbf{S}_t^{\mathcal{A}}(w)) \leq \mathcal{A}(\mathbf{S}_s^{\mathcal{A}}(w)) \quad \forall 0 \leq s \leq t$$

for all  $w \in \mathcal{P}$ , is called  $\kappa$ -flow or *gradient flow* of  $\mathcal{A}$ .

Notice that [7, 14] if  $\mathcal{A}$  induces a  $\kappa$ -flow on  $(\mathcal{P}, \mathbf{d})$  for some  $\kappa \in \mathbb{R}$ , then  $\mathcal{A}$  is  $\kappa$ -convex along geodesics in  $(\mathcal{P}, \mathbf{d})$  in the sense of McCann [20].

**Theorem 2.2** (Flow interchange lemma [19, Thm. 3.2]). *Let  $\mathcal{B}$  be a proper and lower semicontinuous functional on  $(\mathcal{P}, \mathbf{d})$  and assume that there exists a  $\lambda$ -flow  $\mathbf{S}^{\mathcal{B}}$  for some  $\lambda \in \mathbb{R}$ . Let furthermore  $\mathcal{A}$  be another proper, lower semicontinuous functional on  $(\mathcal{P}, \mathbf{d})$  such that  $\operatorname{Dom}(\mathcal{A}) \subset \operatorname{Dom}(\mathcal{B})$ . Assume that, if  $\tau > 0$  and  $\tilde{w} \in \mathcal{P}$  are such that the functional  $u \mapsto \frac{1}{2\tau} \mathbf{d}^2(u, \tilde{w}) + \mathcal{A}(u)$  is proper, it also possesses a minimizer  $w$  on  $\mathcal{P}$ . Then, the following holds:*

$$\mathcal{B}(w) + \tau \mathbf{D}^{\mathcal{B}} \mathcal{A}(w) + \frac{\lambda}{2} \mathbf{d}^2(w, \tilde{w}) \leq \mathcal{B}(\tilde{w}).$$

There,  $\mathbf{D}^{\mathcal{B}} \mathcal{A}(w)$  denotes the dissipation of the functional  $\mathcal{A}$  along the  $\lambda$ -flow  $\mathbf{S}^{\mathcal{B}}$  of the functional  $\mathcal{B}$ , i.e.

$$\mathbf{D}^{\mathcal{B}} \mathcal{A}(w) := \limsup_{h \searrow 0} \frac{\mathcal{A}(w) - \mathcal{A}(\mathbf{S}_h^{\mathcal{B}}(w))}{h}.$$

With the uniform estimates on  $u_\tau$  derived with the flow interchange lemma, one seeks to pass to the limit  $\tau \searrow 0$  w.r.t. a suitably strong notion of convergence. To this end, the following extension of the classical Aubin-Lions compactness lemma to the metric setting is useful.

**Theorem 2.3** (Extension of the Aubin-Lions lemma [24, Thm. 2]). *Let  $\mathbf{Y}$  be a Banach space and let  $\mathcal{A} : \mathbf{Y} \rightarrow [0, \infty]$  be lower semicontinuous and have relatively compact sublevels in  $\mathbf{Y}$ . Let furthermore  $\mathbf{W} : \mathbf{Y} \times \mathbf{Y} \rightarrow [0, \infty]$  be lower semicontinuous and such that  $\mathbf{W}(u, \tilde{u}) = 0$  for  $u, \tilde{u} \in \text{Dom}(\mathcal{A})$  implies  $u = \tilde{u}$ .*

*If for a sequence  $(U_k)_{k \in \mathbb{N}}$  of measurable functions  $U_k : (0, T) \rightarrow \mathbf{Y}$ , one has*

$$\sup_{k \in \mathbb{N}} \int_0^T \mathcal{A}(U_k(t)) dt < \infty \quad \text{and} \quad (2.1)$$

$$\limsup_{h \searrow 0} \sup_{k \in \mathbb{N}} \int_0^{T-h} \mathbf{W}(U_k(t+h), U_k(t)) dt = 0, \quad (2.2)$$

*then there exists a subsequence that converges in measure w.r.t.  $t \in (0, T)$  to a limit  $U : (0, T) \rightarrow \mathbf{Y}$ .*

We conclude this preliminary section with some elementary properties of the mobility functions  $m$  and their induced distances.

**Proposition 2.4** (Properties of  $m$  and  $\mathbf{W}_{m(t, \cdot)}$ ). *The following statements hold:*

- (a) *For each fixed  $t \geq 0$ , the distance functional  $\mathbf{W}_{m(t, \cdot)}$  is in both arguments lower semicontinuous with respect to weak\*-convergence in  $\mathcal{P}$ .*
- (b) *For all  $0 \leq t_0 \leq t_1$ , one has*

$$\mathbf{X}(t_0) \subset \mathbf{X}(t_1) \subset \mathbf{X}.$$

- (c) *For all  $T > 0$ , all  $t \in [0, T]$  and all  $z \in [0, S(t))$ , one has, for  $C(T) := \max_{t \in [0, T]} \lim_{z' \searrow 0} \partial_z m(t, z') < \infty$ :*

$$\begin{aligned} \partial_z m(t, z) &\leq C(T), \\ m(t, z) &\leq C(T)z =: \overline{m}(T, z). \end{aligned}$$

- (d) *At each  $T > 0$ ,  $\mathbf{W}_{\overline{m}(T, \cdot)}$  is a scalar multiple of the  $L^2$ -Wasserstein distance on the space of probability measures and the following estimate holds for  $t \in [0, T]$  and  $u, \tilde{u} \in \mathbf{X}(t)$ :*

$$\mathbf{W}_2(u, \tilde{u}) \leq \sqrt{C(T)} \mathbf{W}_{\overline{m}(T, \cdot)}(u, \tilde{u}) \leq \sqrt{C(T)} \mathbf{W}_{m(t, \cdot)}(u, \tilde{u}).$$

*Proof.* The lower semicontinuity of  $\mathbf{W}_{m(t, \cdot)}$  is classical, see [8, 16]. Obviously, (b) holds as  $S$  is nondecreasing (M1). The estimates in (c) are a straightforward consequence of (M1)–(M3), since  $\partial_z m(t, \cdot)$  is nonincreasing. Finally, (d) can easily be derived using (c) and the characterization of the  $L^2$ -Wasserstein distance by the Benamou-Brenier formula.  $\square$

### 3. THE VARIATIONAL SCHEME

In this section, we prove the well-posedness of the scheme from Definition 1.2, i.e., the sequence of successive minimizers  $(u_\tau^n)_{n \in \mathbb{N}}$  in (1.10) exists, and investigate certain regularity properties. First, some elementary properties of the free energy  $\mathcal{E}$  are summarized.

**Lemma 3.1** (Properties of  $\mathcal{E}$ , case (E1)). *The following statements hold:*

- (a) *For all  $u \in \mathbf{X} \cap L^2(\mathbb{R}^d)$ , one has*

$$-\infty < \inf_{x \in \mathbb{R}^d} \phi \leq \frac{\gamma_0}{2} \|u\|_{L^2}^2 + \inf_{x \in \mathbb{R}^d} \phi \leq \mathcal{E}(u) \leq \frac{\gamma_1}{2} \|u\|_{L^2}^2 + C_\phi(\mathbf{m}_2(u) + 1) < \infty,$$

*where  $C_\phi > 0$  is such that  $\phi(x) \leq C_\phi(|x|^2 + 1)$  for all  $x \in \mathbb{R}^d$ .*

- (b) *Let a sequence  $(u_k)_{k \in \mathbb{N}}$  in  $\mathbf{X} \cap L^2(\mathbb{R}^d)$  be given and assume that  $u_k$  converges weakly\* in  $\mathcal{P}$  and weakly in the space  $L^2(\mathbb{R}^d)$  to some  $u \in \mathbf{X} \cap L^2(\mathbb{R}^d)$ . Then,*

$$\mathcal{E}(u) \leq \liminf_{k \rightarrow \infty} \mathcal{E}(u_k) < \infty.$$



*Proof.* By Taylor's theorem, we have  $f(z) = f(z) + zf'(0) + \frac{1}{2}f''(\tilde{z})z^2$  for all  $z$  and some  $\tilde{z} \in (0, z)$ . Hence, using the bounds on  $f''$  from (E1) and the estimate  $-\infty < \inf \phi \leq \phi(x) \leq C_\phi(|x|^2 + 1) < \infty$ , the claim in (a) follows immediately. Notice in particular that  $f$  is a nonnegative and convex function. Therefore, the map  $u \mapsto \int_{\mathbb{R}^d} f(u(x)) dx$  is weakly lower semicontinuous in  $L^2(\mathbb{R}^d)$ . Together with the—as  $\phi$  is bounded from below—obvious weak\*-lower semicontinuity of  $u \mapsto \int_{\mathbb{R}^d} \phi(x)u(x) dx$  [1, Lemma 5.1.7], part (b) follows.  $\square$

Similar arguments show an analogous result in the case of gradient-dependent energy of the form in (E2):

**Lemma 3.2** (Properties of  $\mathcal{E}$ , case (E2)). *The following statements hold:*

(a) *For all  $u \in \mathbf{X} \cap H^1(\mathbb{R}^d)$ , one has*

$$-\infty < \inf_{x \in \mathbb{R}^d} \phi \leq \frac{\gamma_0}{2} \|u\|_{H^1}^2 + \inf_{x \in \mathbb{R}^d} \phi \leq \mathcal{E}(u) \leq \frac{\gamma_1}{2} \|u\|_{H^1}^2 + C_\phi(\mathbf{m}_2(u) + 1) < \infty,$$

*where  $C_\phi > 0$  is such that  $\phi(x) \leq C_\phi(|x|^2 + 1)$  for all  $x \in \mathbb{R}^d$ .*

(b) *Let a sequence  $(u_k)_{k \in \mathbb{N}}$  in  $\mathbf{X} \cap H^1(\mathbb{R}^d)$  with uniformly bounded second moments  $\mathbf{m}_2(u_k)$  be given and assume that  $u_k$  converges weakly\* in  $\mathcal{P}$  and weakly in the space  $H^1(\mathbb{R}^d)$  to some  $u \in \mathbf{X} \cap H^1(\mathbb{R}^d)$ . Then,*

$$\mathcal{E}(u) \leq \liminf_{k \rightarrow \infty} \mathcal{E}(u_k) < \infty.$$

**3.1. Well-posedness and classical properties.** This paragraph is concerned with the well-posedness of the scheme (1.10) and properties resembling the classical estimates known for the autonomous case.

**Proposition 3.3** (Well-posedness and classical estimates, case (E1)). *Assume that  $u_0 \in \mathbf{X}(0) \cap L^2(\mathbb{R}^d)$  and let  $\bar{\tau} > 0$ . Then, for all  $\tau \in (0, \bar{\tau}]$  and all  $n \in \mathbb{N}$ , the map*

$$\mathcal{E}_n(u) := \frac{1}{2\tau} \mathbf{W}_{m(n\tau, \cdot)}(u_\tau^{n-1}, u)^2 + \mathcal{E}(u) \quad (3.1)$$

*possesses a minimizer  $u_\tau^n$  on  $\mathbf{X}$  belonging to  $\mathbf{X}(n\tau) \cap L^2(\mathbb{R}^d)$ . Furthermore, the following estimates hold:*

$$\mathcal{E}(u_\tau^n) \leq \mathcal{E}(u_\tau^{n-1}) \leq \mathcal{E}(u_0) < \infty, \quad \text{for all } n \in \mathbb{N}, \quad (3.2)$$

$$\sum_{n=1}^{\infty} \mathbf{W}_{m(n\tau, \cdot)}(u_\tau^n, u_\tau^{n-1})^2 \leq 2\tau(\mathcal{E}(u_0) - \inf_{u \in \mathbf{X}} \mathcal{E}(u)) < \infty, \quad (3.3)$$

$$\mathbf{W}_2(u_\tau(t), u_\tau(s)) \leq \sqrt{2C(T + \bar{\tau})(\mathcal{E}(u_0) - \inf_{u \in \mathbf{X}} \mathcal{E}(u)) \max(\tau, |t - s|)}, \quad (3.4)$$

*for all  $s, t \in [0, T]$  and all  $T > 0$ .*

*Proof.* We proceed by induction on  $n \in \mathbb{N}$  and consider a minimizing sequence  $(u_k)_{k \in \mathbb{N}}$  for  $\mathcal{E}_n$  which consequently is a sequence in  $\mathbf{X}(n\tau) \cap L^2(\mathbb{R}^d)$ . Using the bounds from Proposition 2.4(b), one has for a suitable constant  $C > 0$  that  $\|u_k\|_{L^2} \leq C$  and  $\mathbf{m}_2(u_k) \leq C$  for all  $k \in \mathbb{N}$ . Subsequently, the Banach-Alaoglu and Prokhorov theorems yield the existence of a (non-relabelled) subsequence and a limit  $u \in \mathbf{X}(n\tau) \cap L^2(\mathbb{R}^d)$  such that  $u_k$  converges to  $u$  both weakly\* in  $\mathcal{P}$  and weakly in  $L^2(\mathbb{R}^d)$ . Since  $\mathcal{E}$  and  $\mathbf{W}_{m(n\tau, \cdot)}(u_\tau^{k-1}, \cdot)$  are lower semicontinuous w.r.t. these convergences, the limit  $u$  indeed is a minimizer of  $\mathcal{E}_n$ .

The minimizing property of  $u_\tau^n$  immediately yields (3.2), and after summing up, also (3.3). For the Hölder type estimate (3.4), we fix  $T > 0$  and  $s, t \in [0, T]$  and find, using the triangle inequality and Proposition 2.4(d):

$$\mathbf{W}_2(u_\tau(t), u_\tau(s)) \leq \sum_{n=M}^N \mathbf{W}_2(u_\tau^n, u_\tau^{n-1}) \leq \sqrt{C(T + \bar{\tau})} \sum_{n=M}^N \mathbf{W}_{m(n\tau, \cdot)}(u_\tau^n, u_\tau^{n-1})$$



for some  $M, N \in \mathbb{N}$  with  $M \leq N$ . We proceed using Hölder's inequality and (3.3) to obtain the desired estimate:

$$\begin{aligned} \sqrt{C(T+\bar{\tau})} \sum_{n=M}^N \mathbf{W}_{m(n\tau, \cdot)}(u_\tau^n, u_\tau^{n-1}) &\leq \left( C(T+\bar{\tau})(N-M+1) \sum_{n=M}^N \mathbf{W}_{m(n\tau, \cdot)}(u_\tau^n, u_\tau^{n-1})^2 \right)^{1/2} \\ &\leq \sqrt{2C(T+\bar{\tau})(\mathcal{E}(u_0) - \inf_{u \in \mathbf{X}} \mathcal{E}(u)) \max(\tau, |t-s|)}. \quad \square \end{aligned}$$

By an easy adaptation of the proof, one obtains

**Proposition 3.4** (Well-posedness and classical estimates, case (E2)). *Assume that  $u_0 \in \mathbf{X}(0) \cap H^1(\mathbb{R}^d)$  and let  $\bar{\tau} > 0$ . Then, for all  $\tau \in (0, \bar{\tau}]$  and all  $n \in \mathbb{N}$ , the map from (3.1) possesses a minimizer  $u_\tau^n$  on  $\mathbf{X}$  belonging to  $\mathbf{X}(n\tau) \cap H^1(\mathbb{R}^d)$ , and the estimates (3.2)–(3.4) hold.*

**3.2. Additional regularity.** The minimizers of  $\mathcal{E}_n$  enjoy a certain regularity property which is crucial for the passage to the continuous-time limit afterwards. To this end, we introduce for  $t \geq 0$  the *time-dependent heat entropy*

$$\mathcal{H}_{(\cdot)} : \mathbb{R}_{\geq 0} \times \mathbf{X} \rightarrow \mathbb{R}_{\infty}, \quad \mathcal{H}_t(u) := \begin{cases} \int_{\mathbb{R}^d} h(t, u(t, x)) \, dx & \text{if } u \in \mathbf{X}(t), \\ +\infty & \text{otherwise.} \end{cases} \quad (3.5)$$

We first prove some elementary properties of  $\mathcal{H}$ :

**Lemma 3.5** (Time-dependent heat entropy). *The following statements hold :*

(a) *There exists a continuous function  $C : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  such that for each fixed  $t \geq 0$  and all  $u \in \mathbf{X}(t) \cap L^2(\mathbb{R}^d)$ :*

$$|\mathcal{H}_t(u)| \leq C(t)(\|u\|_{L^2}^2 + \mathbf{m}_2(u) + 1) < \infty. \quad (3.6)$$

(b) *For each fixed  $t \geq 0$ ,  $\mathcal{H}_t$  induces a 0-flow  $\mathcal{S}^{\mathcal{H}_t}$  on  $\mathbf{X}(t)$  coinciding with the heat flow, viz.*

$$\partial_s \mathcal{S}_s^{\mathcal{H}_t}(u) = \Delta \mathcal{S}_s^{\mathcal{H}_t}(u), \quad \mathcal{S}_0^{\mathcal{H}_t}(u) = u, \quad \text{for } u \in \mathbf{X}(t).$$

*Proof.* Thanks to the bounds on  $h$  from (M4), one has for all  $t \geq 0$  and all  $u \in \mathbf{X}(t)$  that

$$\mathcal{H}_t(u) \leq H_1(t) \int_{\mathbb{R}^d} u^{\alpha_1(t)} \, dx + H_2(t) \int_{\mathbb{R}^d} u^{\alpha_2(t)+1} \, dx.$$

The first integral can be estimated using Hölder's inequality (recall that  $\alpha_1(t) < 1$ ):

$$\begin{aligned} \int_{\mathbb{R}^d} u^{\alpha_1(t)} \, dx &= \int_{\mathbb{R}^d} (u(x)(|x|^2 + 1))^{\alpha_1(t)} (|x|^2 + 1)^{-\alpha_1(t)} \, dx \\ &\leq (\mathbf{m}_2(u) + 1)^{\alpha_1(t)} \left( \int_{\mathbb{R}^d} (|x|^2 + 1)^{-\frac{\alpha_1(t)}{1-\alpha_1(t)}} \, dx \right)^{1-\alpha_1(t)}, \end{aligned}$$

and  $\int_{\mathbb{R}^d} (|x|^2 + 1)^{-\frac{\alpha_1(t)}{1-\alpha_1(t)}} \, dx$  is finite since  $\frac{-2\alpha_1(t)}{1-\alpha_1(t)} + d - 1 < -1$  thanks to  $\alpha_1(t) > \frac{d}{d+2}$ . The second integral above can be controlled with the  $L^p$  interpolation inequality as follows:

$$\int_{\mathbb{R}^d} u^{\alpha_2(t)+1} \, dx \leq \|u\|_{L^1}^{1-\theta(t)} \|u\|_{L^2}^{\theta(t)} = \|u\|_{L^2}^{\theta(t)},$$

for a suitable  $\theta(t) \in (0, 1)$  depending continuously on  $t$ . All in all, applying Young's inequality provides (3.6).

Since  $m(t, \cdot)$  is induced by  $h(t, \cdot)$  and since  $\mathcal{H}_t$  is finite on  $\mathbf{X}(t)$ , a 0-flow on  $\mathbf{X}(t)$  is induced by  $\mathcal{H}_t$ , for each fixed  $t \geq 0$  (see [17, 29]). By construction,  $\mathcal{S}^{\mathcal{H}_t}$  is the heat flow.  $\square$

**Proposition 3.6** (Additional regularity, case (E1)). *Assume that  $u_0 \in \mathbf{X}(0) \cap L^2(\mathbb{R}^d)$  and let  $\bar{\tau} > 0$ . Then, there exists  $C > 0$  such that for all  $\tau \in (0, \bar{\tau}]$  and all  $n \in \mathbb{N}$ , one has*

$$\tau \|\nabla u_\tau^n\|_{L^2}^2 \leq C(\mathcal{H}_{n\tau}(u_\tau^{n-1}) - \mathcal{H}_{n\tau}(u_\tau^n) + \tau). \quad (3.7)$$

*Proof.* Our proof is a (by now almost classical) application of the *flow interchange* technique from [19]. The suitable auxiliary flow is the aforementioned 0-flow  $\mathcal{S}^{\mathcal{H}_{n\tau}}$  of the functional  $\mathcal{H}_t$  at  $t = n\tau$ , for each fixed  $n \in \mathbb{N}$ . We calculate the dissipation of  $\mathcal{E}$  along the semigroup  $(\mathcal{S}_s^{\mathcal{H}_{n\tau}}(u_\tau^n))_{s \geq 0}$  and write  $u_s := \mathcal{S}_s^{\mathcal{H}_{n\tau}}(u_\tau^n)$  for brevity:

$$-\frac{d}{ds}\mathcal{E}(u_s) = - \int_{\mathbb{R}^d} (f'(u_s) + \phi) \Delta u_s \, dx.$$

Integrating by parts and using (E1), we obtain

$$-\frac{d}{ds}\mathcal{E}(u_s) = \int_{\mathbb{R}^d} (f''(u_s)|\nabla u_s|^2 - u_s \Delta \phi) \, dx \geq \gamma_0 \int_{\mathbb{R}^d} |\nabla u_s|^2 \, dx - \|\Delta \phi\|_{L^\infty}.$$

As  $s \searrow 0$ , one has

$$D^{\mathcal{H}_{n\tau}} \mathcal{E}(u_\tau^n) \geq \liminf_{s \searrow 0} \left( -\frac{d}{ds} \mathcal{E}(u_s) \right) \geq \gamma_0 \int_{\mathbb{R}^d} |\nabla u_\tau^n|^2 \, dx - \|\Delta \phi\|_{L^\infty}.$$

We apply the flow interchange lemma (Theorem 2.2) and rearrange to obtain the desired result:

$$\mathcal{H}_{n\tau}(u_\tau^n) + \tau \left[ \gamma_0 \int_{\mathbb{R}^d} |\nabla u_s|^2 \, dx - \|\Delta \phi\|_{L^\infty} \right] \leq \mathcal{H}_{n\tau}(u_\tau^{n-1}). \quad \square$$

A similar result is also true for gradient-dependent energy:

**Proposition 3.7** (Additional regularity, case (E2)). *Assume that  $u_0 \in \mathbf{X}(0) \cap H^1(\mathbb{R}^d)$  and let  $\bar{\tau} > 0$ . Then, there exists  $C > 0$  such that for all  $\tau \in (0, \bar{\tau}]$  and all  $n \in \mathbb{N}$ , one has*

$$\tau \|\nabla^2 u_\tau^n\|_{L^2}^2 \leq C(\mathcal{H}_{n\tau}(u_\tau^{n-1}) - \mathcal{H}_{n\tau}(u_\tau^n) + \tau). \quad (3.8)$$

*Proof.* As in the proof of Proposition 3.6, we derive the dissipation of  $\mathcal{E}$  along  $(\mathcal{S}_s^{\mathcal{H}_{n\tau}}(u_\tau^n))_{s \geq 0}$ , integrate by parts and use (E2):

$$\begin{aligned} -\frac{d}{ds}\mathcal{E}(u_s) &= - \int_{\mathbb{R}^d} [-\operatorname{div} \nabla_p f(\nabla u_s, u_s) + \partial_z f(\nabla u_s, u_s) + \phi] \Delta u_s \, dx \\ &= - \int_{\mathbb{R}^d} \left[ \sum_{i=1}^d (-\operatorname{div} \nabla_p f(\nabla u_s, u_s) + \partial_z f(\nabla u_s, u_s)) \partial_{x_i} \partial_{x_i} u_s + u_s \Delta \phi \right] \, dx \\ &= - \int_{\mathbb{R}^d} \left[ \sum_{i=1}^d \nabla_p f(\nabla u_s, u_s) \cdot \partial_{x_i} \partial_{x_i} \nabla u_s + \partial_z f(\nabla u_s, u_s) \partial_{x_i} \partial_{x_i} u_s + u_s \Delta \phi \right] \, dx \\ &= \int_{\mathbb{R}^d} \left[ \sum_{i=1}^d \partial_{x_i} \nabla_p f(\nabla u_s, u_s) \cdot \partial_{x_i} \nabla u_s + \partial_{x_i} \partial_z f(\nabla u_s, u_s) \partial_{x_i} u_s - u_s \Delta \phi \right] \, dx \\ &= \int_{\mathbb{R}^d} \left[ \sum_{i=1}^d \begin{pmatrix} \partial_{x_i} \nabla u_s \\ \partial_{x_i} u_s \end{pmatrix}^\top \nabla_{(p,z)}^2 f(\nabla u_s, u_s) \begin{pmatrix} \partial_{x_i} \nabla u_s \\ \partial_{x_i} u_s \end{pmatrix} - u_s \Delta \phi \right] \, dx \\ &\geq \int_{\mathbb{R}^d} \left[ \gamma_0 \sum_{i=1}^d ((\partial_{x_i} u_s)^2 + |\partial_{x_i} \nabla u_s|^2) - u_s \Delta \phi \right] \, dx \\ &\geq \gamma_0 \|\nabla^2 u_s\|_{L^2}^2 - \|\Delta \phi\|_{L^\infty}. \end{aligned}$$

From there, the proof goes along the same lines as the proof of Proposition 3.6.  $\square$

We can now summarize the relevant *a priori* estimates on the discrete solution  $u_\tau$ .

**Lemma 3.8** (*A priori* estimates, case (E1)). *Let  $u_0 \in \mathbf{X}(0) \cap L^2(\mathbb{R}^d)$ ,  $\bar{\tau} > 0$  and  $T > 0$  be given. Then, there exists a constant  $C > 0$  such that for all  $\tau \in (0, \bar{\tau}]$ :*

$$(a) \quad \|u_\tau\|_{L^\infty([0,T];L^2)} \leq C,$$

- (b)  $\sup_{t \in [0, T]} \mathbf{m}_2(u_\tau(t, \cdot)) \leq C,$   
(c)  $\|u_\tau\|_{L^2([0, T]; H^1)} \leq C.$

*Proof.* The parts (a) and (b) are direct consequences of the estimates (3.2) and (3.4), respectively. For part (c), we use (3.7) to see for  $N := \lfloor \frac{T}{\tau} \rfloor + 1$  that

$$\|\nabla u_\tau\|_{L^2([0, T]; L^2)}^2 \leq \tau \sum_{n=1}^N \|\nabla u_\tau^n\|_{L^2}^2 \leq C \sum_{n=1}^N (\mathcal{H}_{n\tau}(u_\tau^{n-1}) - \mathcal{H}_{n\tau}(u_\tau^n) + \tau). \quad (3.9)$$

Inserting a suitable term and employing the local Lipschitz condition for  $h$  from (M4) yields for a fixed  $n$ :

$$\begin{aligned} \mathcal{H}_{n\tau}(u_\tau^{n-1}) - \mathcal{H}_{n\tau}(u_\tau^n) &= \mathcal{H}_{n\tau}(u_\tau^{n-1}) - \mathcal{H}_{(n-1)\tau}(u_\tau^{n-1}) + \mathcal{H}_{(n-1)\tau}(u_\tau^{n-1}) - \mathcal{H}_{n\tau}(u_\tau^n) \\ &\leq \tau \int_{\mathbb{R}^d} \left[ L_1(T + \bar{\tau})(u_\tau^{n-1})^{\beta_1(T + \bar{\tau})} + L_2(T + \bar{\tau})(u_\tau^{n-1})^{\beta_2(T + \bar{\tau}) + 1} \right] dx + \mathcal{H}_{(n-1)\tau}(u_\tau^{n-1}) - \mathcal{H}_{n\tau}(u_\tau^n). \end{aligned}$$

Thus, we get from (3.9) simplifying the telescopic sum that

$$\begin{aligned} \|\nabla u_\tau\|_{L^2([0, T]; L^2)}^2 &\leq \tau \sum_{n=1}^N C \left[ 1 + \int_{\mathbb{R}^d} \left[ L_1(T + \bar{\tau})(u_\tau^{n-1})^{\beta_1(T + \bar{\tau})} + L_2(T + \bar{\tau})(u_\tau^{n-1})^{\beta_2(T + \bar{\tau}) + 1} \right] dx \right] \\ &\quad + C(\mathcal{H}_0(u_0) - \mathcal{H}_{N\tau}(u_\tau^N)). \end{aligned}$$

Proceeding as in the proof of (3.6) from Lemma 3.5 gives us for a constant  $C' > 0$  that

$$\|\nabla u_\tau\|_{L^2([0, T]; L^2)}^2 \leq C(T + \bar{\tau}) + C'\tau \sum_{n=0}^N [\mathbf{m}_2(u_\tau^n) + \|u_\tau^n\|_{L^2} + 1],$$

which can be controlled by a finite constant independent of  $\tau$  with the parts (a) and (b) of this proposition.  $\square$

Analogously, one shows:

**Lemma 3.9** (*A priori estimates, case (E2)*). *Let  $u_0 \in \mathbf{X}(0) \cap H^1(\mathbb{R}^d)$ ,  $\bar{\tau} > 0$  and  $T > 0$  be given. Then, there exists a constant  $C > 0$  such that for all  $\tau \in (0, \bar{\tau}]$ :*

- (a)  $\|u_\tau\|_{L^\infty([0, T]; H^1)} \leq C,$   
(b)  $\sup_{t \in [0, T]} \mathbf{m}_2(u_\tau(t, \cdot)) \leq C,$   
(c)  $\|u_\tau\|_{L^2([0, T]; H^2)} \leq C.$

#### 4. WEAK FORMULATION AND PASSAGE TO THE CONTINUOUS-TIME LIMIT

First, we derive an approximate weak formulation of equation (1.1) corresponding to the Euler-Lagrange equation associated to the minimization problem in (1.10). Afterwards, we show that  $u_\tau$  converges in a suitable sense to a map  $u$  which fulfills the (time-continuous) distributional formulation of equation (1.1).

**4.1. Discrete weak formulation.** In this paragraph, we derive an approximate weak formulation for equation (1.1) in discrete time. Again, our method relies on the flow interchange lemma (Theorem 2.2). Here, the auxiliary flow is the  $\lambda$ -flow of the time-dependent *regularized potential energy*

$$\mathcal{V}_{(\cdot)} : \mathbb{R}_{\geq 0} \times \mathbf{X} \rightarrow \mathbb{R}_\infty, \quad \mathcal{V}_t(u) = \sqrt{\tau} \mathcal{H}_t(u) + \int_{\mathbb{R}^d} u \eta dx,$$

for some test function  $\eta \in C_c^\infty(\mathbb{R}^d)$  and some  $\tau > 0$ :

**Lemma 4.1** (*t*-uniform  $\lambda$ -convexity of  $\mathcal{V}_t$ ). *Let  $T > 0$ ,  $\eta \in C_c^\infty(\mathbb{R}^d)$ ,  $\bar{\tau} > 0$  and  $\tau \in (0, \bar{\tau}]$ . There exists a constant  $C > 0$  such that for every  $t \in [0, T]$  and with  $\lambda := -\frac{C}{\sqrt{\tau}}$ , the functional  $\mathcal{V}_t$  is  $\lambda$ -convex and induces a  $\lambda$ -flow  $S_s^{\mathcal{V}_t}$  on  $\mathbf{X}$ . The flow  $S_s^{\mathcal{V}_t}$  solves the viscous continuity equation, viz.*

$$\partial_s S_s^{\mathcal{V}_t}(u) = \sqrt{\tau} \Delta S_s^{\mathcal{V}_t}(u) + \operatorname{div}(m(t, S_s^{\mathcal{V}_t}(u)) \nabla \eta), \quad S_0^{\mathcal{V}_t}(u) = u.$$

*Proof.* According to the criterion from [14] (see also [17, 29]), one has to show that for each fixed  $t \in [0, T]$ , all  $z \in (0, S(t))$ ,  $\zeta \in \mathbb{R}$  and all  $w_1, w_2 \in \mathbb{R}$  with  $|w_1| \leq \|\eta\|_{C^2}$ ,  $|w_2| \leq \|\eta\|_{C^2}$ , the following holds for some  $\lambda \in \mathbb{R}$ :

$$-\frac{\sqrt{\tau}}{2} \partial_z^2 m(t, z) \zeta^2 + \frac{1}{2} \partial_z^2 m(t, z) m(t, z) w_1 \zeta + \partial_z m(t, z) m(t, z) w_2 \geq \lambda m(t, z). \quad (4.1)$$

Using Young's inequality, we get

$$\begin{aligned} & -\frac{\sqrt{\tau}}{2} \partial_z^2 m(t, z) \zeta^2 + \frac{1}{2} \partial_z^2 m(t, z) m(t, z) w_1 \zeta + \partial_z m(t, z) m(t, z) w_2 \\ & \geq \left( \frac{1}{8\sqrt{\tau}} \partial_z^2 m(t, z) m(t, z) w_1^2 + \partial_z m(t, z) w_2 \right) m(t, z) \\ & \geq -m(t, z) \left( \frac{1}{8\sqrt{\tau}} M_2(t) w_1^2 + M_1(t) w_2 \right), \end{aligned}$$

with condition (M3) used in the last step. The bounds on  $w_1$ ,  $w_2$  and  $\tau$  and the continuity of  $M_1$  and  $M_2$  yield

$$-m(t, z) \left( \frac{1}{8\sqrt{\tau}} M_2(t) w_1^2 + M_1(t) w_2 \right) \geq -\frac{m(t, z)}{\sqrt{\tau}} \left( \frac{1}{8} R^2 \max_{t \in [0, T]} M_2(t) + \sqrt{\tau} R \max_{t \in [0, T]} M_1(t) \right),$$

so (4.1) holds with  $\lambda := -\frac{1}{8} R^2 \max_{t \in [0, T]} M_2(t) - \sqrt{\tau} R \max_{t \in [0, T]} M_1(t)$ .  $\square$

**Lemma 4.2** (Discrete weak formulation, case (E1)). *Let  $u_0 \in \mathbf{X}(0) \cap L^2(\mathbb{R}^d)$ ,  $\bar{\tau} > 0$  and let test functions  $\eta \in C_c^\infty(\mathbb{R}^d)$  and  $\psi \in C_c^\infty(\mathbb{R}^d) \cap C(\mathbb{R}_{\geq 0})$  be given. Then, there exists a constant  $C > 0$  such that for all  $\tau \in (0, \bar{\tau}]$ , one has*

$$\begin{aligned} & \left| \int_0^\infty \int_{\mathbb{R}^d} \left[ \frac{\psi_\tau(t) - \psi_\tau(t + \tau)}{\tau} u_\tau(t, x) \eta(x) \right. \right. \\ & \quad \left. \left. + \psi_\tau(t) m \left( \left[ \frac{t}{\tau} \right] \tau, u_\tau(t, x) \right) [\nabla u_\tau(t, x) f''(u_\tau(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \right] dx dt \right| \leq C \sqrt{\tau}, \end{aligned} \quad (4.2)$$

with the abbreviation  $\psi_\tau(t) := \psi \left( \left[ \frac{t}{\tau} \right] \tau \right)$ .

*Proof.* Fix test functions  $\eta \in C_c^\infty(\mathbb{R}^d)$  and  $\psi \in C_c^\infty(\mathbb{R}^d) \cap C(\mathbb{R}_{\geq 0})$  and let  $T > 0$ ,  $R > 0$  such that  $\operatorname{supp} \psi \subset [0, T]$  and  $\operatorname{supp} \eta \subset \mathbb{B}_R(0)$ . We first consider  $t = n\tau$  for a fixed  $n \in \mathbb{N}$ . In order to apply the flow interchange principle, we calculate the dissipation of  $\mathcal{E}$  along the semigroup  $S_{(\cdot)}^{\mathcal{V}_{n\tau}}(u_\tau^n) =: u_{(\cdot)}$ :

$$\begin{aligned} -\frac{d}{ds} \mathcal{E}(u_s) &= - \int_{\mathbb{R}^d} (f'(u_s) + \phi) (\sqrt{\tau} \Delta u_s + \operatorname{div}(m(n\tau, u_s) \nabla \eta)) dx \\ &= \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_s) |\nabla u_s|^2 - u_s \Delta \phi) dx + \int_{\mathbb{R}^d} (f''(u_s) \nabla u_s + \nabla \phi) \cdot m(n\tau, u_s) \nabla \eta dx, \end{aligned} \quad (4.3)$$

the last step obtained by integration by parts. We now are concerned with the passage to the limit as  $s \searrow 0$ . First, as  $f''(z) \geq \gamma_0 > 0$  for all  $z \geq 0$ , one has

$$\begin{aligned} & \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_s) |\nabla u_s|^2 - u_s \Delta \phi) dx + \int_{\mathbb{R}^d} (f''(u_s) \nabla u_s + \nabla \phi) \cdot m(n\tau, u_s) \nabla \eta dx \\ &= \sqrt{\tau} \int_{\mathbb{R}^d} (|g(u_s)|^2 - u_s \Delta \phi) dx + \int_{\mathbb{R}^d} (\sqrt{f''(u_s)} g(u_s) + \nabla \phi) \cdot m(n\tau, u_s) \nabla \eta dx, \end{aligned}$$

defining the auxiliary map  $g(u(x)) := \sqrt{f''(u(x))} \nabla u(x)$  for  $u \in \mathbf{X} \cap H^1(\mathbb{R}^d)$  and  $x \in \mathbb{R}^d$ . Using Young's inequality, the estimates from Proposition 2.4(b) and  $f''(z) \leq \gamma_1$ , we arrive at

$$\begin{aligned} & -\frac{d}{ds} \mathcal{E}(u_s) \\ & \geq \frac{\sqrt{\tau}}{2} \int_{\mathbb{R}^d} |g(u_s)|^2 dx - \|\Delta \phi\|_{L^\infty} - C'(\|u_s\|_{L^2}^2 + 1) \\ & \geq \frac{\sqrt{\tau}}{2} \int_{\mathbb{R}^d} \gamma_0 |\nabla u_s|^2 dx - \|\Delta \phi\|_{L^\infty} - C'(\|u_s\|_{L^2}^2 + 1), \end{aligned} \quad (4.4)$$

for some constant  $C' > 0$ . By similar arguments, one proves the boundedness of  $\|u_s\|_{L^2}$  for small  $s \geq 0$ :

$$\begin{aligned} & -\frac{d}{ds} \int_{\mathbb{R}^d} u_s^2 dx = 2\sqrt{\tau} \int_{\mathbb{R}^d} |\nabla u_s|^2 dx + 2 \int_{\mathbb{R}^d} m(n\tau, u_s) \nabla u_s \cdot \nabla \eta dx \\ & \geq \sqrt{\tau} \int_{\mathbb{R}^d} |\nabla u_s|^2 dx - C' \int_{\mathbb{R}^d} u_s^2 dx, \end{aligned} \quad (4.5)$$

so Gronwall's lemma yields  $\|u_s\|_{L^2}^2 \leq \|u_\tau^n\|_{L^2}^2 \exp(C' s)$ .

In view of the flow interchange lemma, we have by (4.4) that both  $\|g(u_s)\|_{L^2}$  and  $\|u_s\|_{H^1}$  are uniformly bounded for small  $s \geq 0$ . By Alaoglu's and Rellich's theorems, we infer (on a suitable subsequence) that  $u_s \rightharpoonup u_\tau^n$  in  $H^1(\mathbb{R}^d)$  as well as  $u_s \rightarrow u_\tau^n$  in  $L^2(\mathbb{B}_R(0))$  and  $u_s(x) \rightarrow u_\tau^n(x)$  for a.e.  $x \in \mathbb{R}^d$ . Since  $f''$  is uniformly bounded, the weak limit of  $(g(u_s))_{s \geq 0}$  as  $s \searrow 0$  in  $L^2(\mathbb{R}^d)$  coincides with  $g(u_\tau^n)$ . Furthermore, since

$$\int_{\mathbb{R}^d} \left[ m(n\tau, u_s) (\sqrt{f''(u_s)} + \nabla \phi) \cdot \nabla \eta \right]^2 dx \leq C \int_{\mathbb{B}_R(0)} (u_s^2 + 1) dx,$$

and the integrand on the right-hand side is uniformly integrable thanks to  $L^2$ -convergence, one gets—using the weak convergence of  $g(u_s)$  and the weak\*-convergence of  $u_s$ :

$$\begin{aligned} & \lim_{s \searrow 0} \int_{\mathbb{R}^d} \left[ -\sqrt{\tau} u_s \Delta \phi + \sqrt{f''(u_s)} g(u_s) + \nabla \phi \right] \cdot m(n\tau, u_s) \nabla \eta dx \\ & = \int_{\mathbb{R}^d} \left[ -\sqrt{\tau} u_\tau^n \Delta \phi + \sqrt{f''(u_\tau^n)} g(u_\tau^n) + \nabla \phi \right] \cdot m(n\tau, u_\tau^n) \nabla \eta dx, \end{aligned}$$

by Vitali's convergence theorem. Combining this with the weak lower semicontinuity of the  $L^2$  norm, we arrive at

$$\begin{aligned} D^{\mathcal{V}_{n\tau}} \mathcal{E}(u_\tau^n) & \geq \liminf_{s \searrow 0} \left( -\frac{d}{ds} \mathcal{E}(u_s) \right) \\ & \geq \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_\tau^n) |\nabla u_\tau^n|^2 - u_\tau^n \Delta \phi) dx + \int_{\mathbb{R}^d} (f''(u_\tau^n) \nabla u_\tau^n + \nabla \phi) \cdot m(n\tau, u_\tau^n) \nabla \eta dx. \end{aligned}$$

Now, the flow interchange lemma (Theorem 2.2) yields

$$\begin{aligned} & \mathcal{V}_{n\tau}(u_\tau^n) + \tau \left[ \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_\tau^n) |\nabla u_\tau^n|^2 - u_\tau^n \Delta \phi) dx + \int_{\mathbb{R}^d} (f''(u_\tau^n) \nabla u_\tau^n + \nabla \phi) \cdot m(n\tau, u_\tau^n) \nabla \eta dx \right] \\ & - \frac{C}{2\sqrt{\tau}} \mathbf{W}_{m(n\tau, \cdot)}^2(u_\tau^n, u_\tau^{n-1}) \leq \mathcal{V}_{n\tau}(u_\tau^{n-1}). \end{aligned}$$

Replacing  $\eta$  with  $-\eta$ , we deduce the following chain of inequalities:

$$\begin{aligned} & -\frac{C}{2\sqrt{\tau}} \mathbf{W}_{m(n\tau, \cdot)}^2(u_\tau^n, u_\tau^{n-1}) + \sqrt{\tau} (\mathcal{H}_{n\tau}(u_\tau^n) - \mathcal{H}_{n\tau}(u_\tau^{n-1})) + \tau \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_\tau^n) |\nabla u_\tau^n|^2 - u_\tau^n \Delta \phi) dx \\ & \leq \int_{\mathbb{R}^d} \eta(u_\tau^n - u_\tau^{n-1}) dx + \tau \int_{\mathbb{R}^d} (f''(u_\tau^n) \nabla u_\tau^n + \nabla \phi) \cdot m(n\tau, u_\tau^n) \nabla \eta dx \end{aligned} \quad (4.6)$$

$$\leq \frac{C}{2\sqrt{\tau}} \mathbf{W}_{m(n\tau, \cdot)}^2(u_\tau^n, u_\tau^{n-1}) - \sqrt{\tau} (\mathcal{H}_{n\tau}(u_\tau^n) - \mathcal{H}_{n\tau}(u_\tau^{n-1})) - \tau\sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_\tau^n) |\nabla u_\tau^n|^2 - u_\tau^n \Delta \phi) dx.$$

Consider now the case of nonnegative-valued test functions  $\psi$ . Multiply (4.6) with  $\psi(n\tau)$  and sum up from  $n = 1$  to  $n = \lfloor \frac{T}{\tau} \rfloor \tau + 1 =: N$  to arrive at

$$\begin{aligned} & - \sum_{n=1}^N \psi(n\tau) \frac{C}{2\sqrt{\tau}} \mathbf{W}_{m(n\tau, \cdot)}^2(u_\tau^n, u_\tau^{n-1}) + \sum_{n=1}^N \psi(n\tau) \sqrt{\tau} (\mathcal{H}_{n\tau}(u_\tau^n) - \mathcal{H}_{n\tau}(u_\tau^{n-1})) \\ & + \sum_{n=1}^N \psi(n\tau) \tau \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_\tau^n) |\nabla u_\tau^n|^2 - u_\tau^n \Delta \phi) dx \\ & \leq \tau \sum_{n=0}^{N-1} \frac{\psi(n\tau) - \psi((n+1)\tau)}{\tau} u_\tau^n \eta dx + \tau \sum_{n=1}^N \psi(n\tau) \int_{\mathbb{R}^d} (f''(u_\tau^n) \nabla u_\tau^n + \nabla \phi) \cdot m(n\tau, u_\tau^n) \nabla \eta dx \quad (4.7) \\ & \leq \sum_{n=1}^N \psi(n\tau) \frac{C}{2\sqrt{\tau}} \mathbf{W}_{m(n\tau, \cdot)}^2(u_\tau^n, u_\tau^{n-1}) - \sum_{n=1}^N \psi(n\tau) \sqrt{\tau} (\mathcal{H}_{n\tau}(u_\tau^n) - \mathcal{H}_{n\tau}(u_\tau^{n-1})) \\ & - \sum_{n=1}^N \psi(n\tau) \tau \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_\tau^n) |\nabla u_\tau^n|^2 - u_\tau^n \Delta \phi) dx. \end{aligned}$$

For a sign-changing test function  $\psi$ , we use (4.7) for the positive and negative part, respectively, and subtract the resulting inequalities. Recalling that  $|\psi| = \psi_+ + \psi_-$ , we get

$$\begin{aligned} & \left| \tau \sum_{n=0}^{N-1} \frac{\psi(n\tau) - \psi((n+1)\tau)}{\tau} u_\tau^n \eta dx + \tau \sum_{n=1}^N \psi(n\tau) \int_{\mathbb{R}^d} (f''(u_\tau^n) \nabla u_\tau^n + \nabla \phi) \cdot m(n\tau, u_\tau^n) \nabla \eta dx \right| \\ & \leq \sum_{n=1}^N |\psi(n\tau)| \frac{C}{2\sqrt{\tau}} \mathbf{W}_{m(n\tau, \cdot)}^2(u_\tau^n, u_\tau^{n-1}) + \left| \sum_{n=1}^N |\psi(n\tau)| \sqrt{\tau} (\mathcal{H}_{n\tau}(u_\tau^n) - \mathcal{H}_{n\tau}(u_\tau^{n-1})) \right| \quad (4.8) \\ & + \left| \sum_{n=1}^N |\psi(n\tau)| \tau \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_\tau^n) |\nabla u_\tau^n|^2 - u_\tau^n \Delta \phi) dx \right|. \end{aligned}$$

The left-hand side in (4.8) coincides with the left-hand side of the desired estimate (4.2)—write in spatio-temporal integral form and recall the definition of  $u_\tau$ . It remains to control the right-hand side in (4.8).

First, thanks to (3.3) and Lemma 3.8(c),

$$\begin{aligned} & \sum_{n=1}^N |\psi(n\tau)| \frac{C}{2\sqrt{\tau}} \mathbf{W}_{m(n\tau, \cdot)}^2(u_\tau^n, u_\tau^{n-1}) + \left| \sum_{n=1}^N |\psi(n\tau)| \tau \sqrt{\tau} \int_{\mathbb{R}^d} (f''(u_\tau^n) |\nabla u_\tau^n|^2 - u_\tau^n \Delta \phi) dx \right| \\ & \leq \|\psi\|_{C^0} \sqrt{\tau} \left[ C(\mathcal{E}(u_0) - \inf_{u \in \mathbf{X}} \mathcal{E}(u)) + \gamma_1 \|u_\tau\|_{L^2([0, T+\bar{\tau}]; H^1)}^2 + (T + \bar{\tau}) \|\Delta \phi\|_{L^\infty} \right] \leq C' \sqrt{\tau}, \end{aligned}$$

for a suitable constant  $C' > 0$ . For the second part on the r.h.s. of (4.8), we first insert suitable terms and use the triangle inequality:

$$\begin{aligned} & \left| \sum_{n=1}^N |\psi(n\tau)| (\mathcal{H}_{n\tau}(u_\tau^n) - \mathcal{H}_{n\tau}(u_\tau^{n-1})) \right| \\ & \leq \sum_{n=1}^N ||\psi(n\tau)| - |\psi((n-1)\tau)|| \mathcal{H}_{n\tau}(u_\tau^{n-1}) + \sum_{n=1}^N |\psi((n-1)\tau)| |\mathcal{H}_{n\tau}(u_\tau^{n-1}) - \mathcal{H}_{(n-1)\tau}(u_\tau^{n-1})| \\ & + \left| \sum_{n=1}^N [|\psi((n-1)\tau)| \mathcal{H}_{(n-1)\tau}(u_\tau^{n-1}) - |\psi(n\tau)| \mathcal{H}_{n\tau}(u_\tau^n)] \right|. \end{aligned}$$

From the regularity of  $\psi$  and Lemma 3.5(a), we have

$$\begin{aligned} & \sum_{n=1}^N \left| |\psi(n\tau)| - |\psi((n-1)\tau)| \right| \mathcal{H}_{n\tau}(u_\tau^{n-1}) \\ & \leq \sum_{n=1}^N \tau \|\psi\|_{C^1} C(1 + \mathbf{m}_2(u_\tau^n) + \|u_\tau^n\|_{L^2}^2) \\ & \leq C(T + \bar{\tau}) \|\psi\|_{C^1} \sup_{t \in [0, T + \bar{\tau}]} \left( 1 + \mathbf{m}_2(u_\tau(t)) + \|u_\tau\|_{L^\infty(L^2([0, T + \bar{\tau}]; L^2))}^2 \right), \end{aligned}$$

which is bounded uniformly in  $\tau$  by Lemma 3.8(a)&(b). Second, the local Lipschitz condition for  $\mathcal{H}_t$  in the  $t$  variable from (M4) yields

$$\begin{aligned} & \sum_{n=1}^N |\psi((n-1)\tau)| \left| \mathcal{H}_{n\tau}(u_\tau^{n-1}) - \mathcal{H}_{(n-1)\tau}(u_\tau^{n-1}) \right| \\ & \leq \|\psi\|_{C^0} \tau \sum_{n=1}^N \int_{\mathbb{R}^d} \left( L_1(T + \bar{\tau})(u_\tau^n)^{\beta_1(T + \bar{\tau})} + L_2(T + \bar{\tau})(u_\tau^n)^{\beta_2(T + \bar{\tau})} \right) dx, \end{aligned}$$

which allows us to proceed as in the proof of (3.6) to obtain

$$\sum_{n=1}^N |\psi((n-1)\tau)| \left| \mathcal{H}_{n\tau}(u_\tau^{n-1}) - \mathcal{H}_{(n-1)\tau}(u_\tau^{n-1}) \right| \leq C.$$

Finally, the last sum on the r.h.s. of (4.8) is a telescopic sum and the remaining term can be controlled with (3.6) again:

$$\begin{aligned} & \left| \sum_{n=1}^N \left[ |\psi((n-1)\tau)| \mathcal{H}_{(n-1)\tau}(u_\tau^{n-1}) - |\psi(n\tau)| \mathcal{H}_{n\tau}(u_\tau^n) \right] \right| \\ & = |\psi(N\tau) \mathcal{H}_{N\tau}(u_\tau^N)| \leq C \|\psi\|_{C^0} \sup_{t \in [0, T + \bar{\tau}]} \left( 1 + \mathbf{m}_2(u_\tau(t)) + \|u_\tau\|_{L^\infty(L^2([0, T + \bar{\tau}]; L^2))}^2 \right). \end{aligned}$$

All in all, we obtain the desired estimate (4.2).  $\square$

By the same method, one considers the fourth order case (E2).

**Lemma 4.3** (Discrete weak formulation, case (E2)). *Let  $u_0 \in \mathbf{X}(0) \cap H^1(\mathbb{R}^d)$ ,  $\bar{\tau} > 0$  and let test functions  $\eta \in C_c^\infty(\mathbb{R}^d)$  and  $\psi \in C_c^\infty((0, \infty)) \cap C(\mathbb{R}_{\geq 0})$  be given. Then, there exists a constant  $C > 0$  such that for all  $\tau \in (0, \bar{\tau}]$ , one has*

$$\begin{aligned} & \left| \int_0^\infty \int_{\mathbb{R}^d} \left[ \frac{\psi_\tau(t) - \psi_\tau(t + \tau)}{\tau} u_\tau \eta + \psi_\tau q_\tau \cdot \nabla \phi \right. \right. \\ & \quad \left. \left. + \psi_\tau \operatorname{tr} \left\{ \begin{pmatrix} \mathbb{1}_d \operatorname{div} q_\tau & q_\tau \end{pmatrix} \nabla_{(p,z)}^2 f(\nabla u_\tau, u_\tau) \begin{pmatrix} \nabla^2 u_\tau \\ (\nabla u_\tau)^\top \end{pmatrix} \right\} \right] dx dt \right| \leq C \sqrt{\tau}, \end{aligned} \quad (4.9)$$

with the abbreviations  $q_\tau := m(\lceil \frac{t}{\tau} \rceil \tau, u_\tau) \nabla \eta$  and  $\psi_\tau(t) := \psi(\lceil \frac{t}{\tau} \rceil \tau)$ .

*Proof.* Again, we calculate the dissipation of  $\mathcal{E}$  along the semigroup  $(S_s^{\mathcal{V}_{n\tau}}(u_\tau^n))_{s \geq 0}$ :

$$-\frac{d}{ds} \mathcal{E}(u_s) = - \int_{\mathbb{R}^d} [-\operatorname{div} \nabla_p f(\nabla u_s, u_s) + \partial_z f(\nabla u_s, u_s) + \phi] [\sqrt{\tau} \Delta u_s + \operatorname{div}(m(n\tau, u_s) \nabla \eta)] dx.$$

Write  $q_s := m(n\tau, u_s) \nabla \eta$  for  $s > 0$  and  $q_\tau^n := m(n\tau, u_\tau^n) \nabla \eta$  for brevity. Integrating by parts as in the proof of Proposition 3.6 and by elementary calculations,

$$-\frac{d}{ds} \mathcal{E}(u_s)$$



$$\begin{aligned}
&= \sqrt{\tau} \int_{\mathbb{R}^d} \left[ \sum_{i=1}^d \left( \frac{\partial_{x_i} \nabla u_s}{\partial_{x_i} u_s} \right)^T \nabla_{(p,z)}^2 f(\nabla u_s, u_s) \left( \frac{\partial_{x_i} \nabla u_s}{\partial_{x_i} u_s} \right) - u_s \Delta \phi \right] dx \\
&\quad + \int_{\mathbb{R}^d} [\operatorname{div} q_s \operatorname{div} \nabla_p f(\nabla u_s, u_s) + q_s \cdot (\nabla \partial_z f(\nabla u_s, u_s) + \nabla \phi)] dx \\
&= \sqrt{\tau} \int_{\mathbb{R}^d} \left[ \sum_{i=1}^d \left( \frac{\partial_{x_i} \nabla u_s}{\partial_{x_i} u_s} \right)^T \nabla_{(p,z)}^2 f(\nabla u_s, u_s) \left( \frac{\partial_{x_i} \nabla u_s}{\partial_{x_i} u_s} \right) - u_s \Delta \phi \right] dx \\
&\quad + \int_{\mathbb{R}^d} [\operatorname{div} q_s \operatorname{tr} (\nabla^2 u_s \nabla_p^2 f(\nabla u_s, u_s) + \nabla u_s (\nabla_p \partial_z f(\nabla u_s, u_s))^T) \\
&\quad \quad + q_s \cdot (\partial_z^2 f(\nabla u_s, u_s) \nabla u_s + \nabla^2 u_s \nabla_p \partial_z f(\nabla u_s, u_s) + \nabla \phi)] dx \\
&= \sqrt{\tau} \int_{\mathbb{R}^d} \left[ \sum_{i=1}^d \left( \frac{\partial_{x_i} \nabla u_s}{\partial_{x_i} u_s} \right)^T \nabla_{(p,z)}^2 f(\nabla u_s, u_s) \left( \frac{\partial_{x_i} \nabla u_s}{\partial_{x_i} u_s} \right) - u_s \Delta \phi \right] dx + \int_{\mathbb{R}^d} q_s \cdot \nabla \phi dx \\
&\quad + \int_{\mathbb{R}^d} \operatorname{tr} \left\{ \begin{pmatrix} \mathbb{1}_d \operatorname{div} q_s & q_s \end{pmatrix} \nabla_{(p,z)}^2 f(\nabla u_s, u_s) \begin{pmatrix} \nabla^2 u_s \\ (\nabla u_s)^T \end{pmatrix} \right\} dx.
\end{aligned}$$

Introducing the (uniquely defined) square root  $(\nabla_{(p,z)}^2 f(\nabla u(x), u(x)))^{1/2}$  of the positive definite matrix  $\nabla_{(p,z)}^2 f(\nabla u(x), u(x))$  and the vectors  $g_i(u(x)) := (\nabla_{(p,z)}^2 f(\nabla u(x), u(x)))^{1/2} \begin{pmatrix} \partial_{x_i} \nabla u(x) \\ \partial_{x_i} u(x) \end{pmatrix}$  for  $i \in \{1, \dots, d\}$ ,  $x \in \mathbb{R}^d$  and  $u \in \mathbf{X} \cap H^2(\mathbb{R}^d)$ , the above can also be rewritten as

$$\begin{aligned}
& - \frac{d}{ds} \mathcal{E}(u_s) \\
&= \sqrt{\tau} \int_{\mathbb{R}^d} \left[ \sum_{i=1}^d |g_i(u_s)|^2 - u_s \Delta \phi \right] dx + \int_{\mathbb{R}^d} q_s \cdot \nabla \phi dx \\
&\quad + \sum_{i=1}^d \int_{\mathbb{R}^d} \mathbf{e}_i^T \begin{pmatrix} \mathbb{1}_d \operatorname{div} q_s & q_s \end{pmatrix} (\nabla_{(p,z)}^2 f(\nabla u_s, u_s))^{1/2} g_i(u_s) dx,
\end{aligned}$$

where  $\mathbf{e}_i \in \mathbb{R}^d$  is the  $i^{\text{th}}$  canonical unit vector in  $\mathbb{R}^d$ . Using Young's inequality and the bounds from Proposition 2.4(b), we obtain

$$|q_s| \leq C(T + \bar{\tau}) \|\eta\|_{C^1 u_s}, \quad (4.10)$$

$$|\operatorname{div} q_s| \leq |\partial_z m(t, u_s)| |\nabla u_s| |\nabla \eta| + m(t, u_s) |\Delta \eta| \leq C(T + \bar{\tau}) \|\eta\|_{C^2} (1 + |u_s| + |\nabla u_s|), \quad (4.11)$$

and consequently

$$\begin{aligned}
& - \frac{d}{ds} \mathcal{E}(u_s) \\
&\geq \frac{\sqrt{\tau}}{2} \int_{\mathbb{R}^d} \sum_{i=1}^d |g_i(u_s)|^2 dx - \sqrt{\tau} \|\Delta \phi\|_{L^\infty} - C(\|u_s\|_{L^2}^2 + 1) \\
&\quad - \frac{1}{2} \sum_{i=1}^d \int_{\mathbb{R}^d} \left| \mathbf{e}_i^T \begin{pmatrix} \mathbb{1}_d \operatorname{div} q_s & q_s \end{pmatrix} (\nabla_{(p,z)}^2 f(\nabla u_s, u_s))^{1/2} \right|^2 dx \\
&\geq \frac{\sqrt{\tau}}{2} \int_{\mathbb{R}^d} \sum_{i=1}^d |g_i(u_s)|^2 dx - C'(\|u_s\|_{H^1}^2 + 1) \\
&\geq \frac{\sqrt{\tau}}{2} \gamma_0 \|\nabla^2 u_s\|_{L^2}^2 - C'(\|u_s\|_{H^1}^2 + 1),
\end{aligned}$$

for a suitable constant  $C' > 0$ . Furthermore, by similar arguments, one has

$$\begin{aligned}
& -\frac{d}{ds} \int_{\mathbb{R}^d} |\nabla u_s|^2 dx = \int_{\mathbb{R}^d} [\sqrt{\tau}(\Delta u_s)^2 + \Delta u_s (\partial_z m(n\tau, u_s) \nabla u_s \cdot \nabla \eta + m(n\tau, u_s) \Delta \eta)] dx \\
& \geq \frac{\sqrt{\tau}}{2} \int_{\mathbb{R}^d} (\Delta u_s)^2 dx - \frac{1}{2} \int_{\mathbb{R}^d} (\partial_z m(n\tau, u_s) \nabla u_s \cdot \nabla \eta + m(n\tau, u_s) \Delta \eta)^2 dx \\
& \geq -C \left( \int_{\mathbb{R}^d} |\nabla u_s|^2 dx + \|u_s\|_{L^2} + 1 \right) \\
& \geq -C \int_{\mathbb{R}^d} |\nabla u_s|^2 dx - C',
\end{aligned}$$

where we used in the last step that  $\|u_s\|_{L^2}$  is uniformly bounded for small  $s$ , see (4.5). Hence, Gronwall's lemma yields  $s$ -uniform boundedness of  $u_s$  in  $H^1(\mathbb{R}^d)$ . All in all, we have shown that  $u_s$  is bounded in  $H^2$  as well as that the  $g_i(u_s)$  are bounded in  $L^2(\mathbb{R}^d)$ . From now on, the argumentation for passing to the limit  $s \searrow 0$  and deriving (4.9) is mutatis mutandis the same as in the proof of Lemma 4.2. We omit these technical details here.  $\square$

**4.2. Convergence.** We are now in position to demonstrate the proofs of Theorems 1.5&1.6.

*Proof of Theorem 1.5:* Let a vanishing sequence of step sizes  $(\tau_k)_{k \in \mathbb{N}}$  be given and denote the associated sequence of discrete solutions by  $(u_{\tau_k})_{k \in \mathbb{N}}$ . The uniform estimate on the second moments and the  $L^2([0, T]; H^1(\mathbb{R}^d))$ -norm from Lemma 3.8(b)&(c) yield in combination with the Alaoglu and Prokhorov theorems that  $u_{\tau_k}(t) \rightharpoonup^* u(t)$  in  $\mathcal{P}$  as well as  $u_{\tau_k} \rightharpoonup u$  in  $L^2([0, T]; H^1(\mathbb{R}^d))$ , on a suitable non-relabelled subsequence as  $k \rightarrow \infty$ , for a limit map  $u \in L^2([0, T]; H^1(\mathbb{R}^d))$  with  $u(t) \in \mathbf{X}$  for all  $t \in [0, T]$ . The additional Hölder continuity w.r.t.  $\mathbf{W}_2$  is a consequence of the refined version of the Arzelà-Ascoli theorem [1, Prop. 3.3.1] and (3.4). In order to prove strong convergence, we seek to apply Theorem 2.3: choose  $\mathbf{Y} = L^2(\mathbb{B}_R(0))$  for fixed, but arbitrary  $R > 0$ , the functional  $\mathcal{A} : \mathbf{Y} \rightarrow [0, \infty]$  defined via

$$\mathcal{A}(u) := \begin{cases} \|u\|_{H^1}^2 & \text{if } u \in H^1(\mathbb{B}_R(0)), \\ +\infty & \text{otherwise,} \end{cases}$$

and the pseudo-distance  $\mathbf{W} : \mathbf{Y} \times \mathbf{Y} \rightarrow [0, \infty]$ , defined by

$$\mathbf{W}(u, \tilde{u}) := \inf \{ \mathbf{W}_2(\rho, \tilde{\rho}) : \rho, \tilde{\rho} \in \mathcal{P} \text{ with} \tag{4.12}$$

$$\mathbf{m}_2(\rho) \leq C_{T+\bar{\tau}}, \mathbf{m}_2(\tilde{\rho}) \leq C_{T+\bar{\tau}}, \rho|_{\mathbb{B}_R(0)} = u, \tilde{\rho}|_{\mathbb{B}_R(0)} = \tilde{u} \}, \tag{4.13}$$

where  $C_{T+\bar{\tau}}$  is the constant from Lemma 3.8(b) for  $T+\bar{\tau}$  in place of  $T$ . Rellich's theorem immediately yields relative compactness of sublevels of  $\mathcal{A}$ . Concerning the pseudo-distance  $\mathbf{W}$ , the infimum (if it is finite) in its definition is always attained, since minimizing sequences belong to a weakly\*-compact set by Prokhorov's theorem and the Wasserstein distance is lower semicontinuous w.r.t. weak\*-convergence. One therefore easily shows that  $\mathbf{W}(u, \tilde{u}) = 0$  implies  $u = \tilde{u}$  in  $\mathbf{Y}$ . For the lower semicontinuity property, let a sequence  $(u_k, \tilde{u}_k)_{k \in \mathbb{N}}$  in  $\mathbf{Y} \times \mathbf{Y}$  converging to  $(u, \tilde{u}) \in \mathbf{Y} \times \mathbf{Y}$  be given. Without loss of generality, we may assume that  $\mathbf{W}(u_k, \tilde{u}_k) < \infty$  for all  $k \in \mathbb{N}$ , hence there exists  $(\rho_k, \tilde{\rho}_k)_{k \in \mathbb{N}}$  in  $\mathcal{P} \times \mathcal{P}$  with  $\mathbf{W}(u_k, \tilde{u}_k) = \mathbf{W}_2(\rho_k, \tilde{\rho}_k)$  for all  $k$ . Clearly, on a non-relabelled subsequence,  $\rho_k \rightharpoonup^* \rho'$  and  $\tilde{\rho}_k \rightharpoonup^* \tilde{\rho}'$  in  $\mathcal{P}$ , so

$$\mathbf{W}(u, \tilde{u}) \leq \mathbf{W}_2(\rho', \tilde{\rho}') \leq \liminf_{k \rightarrow \infty} \mathbf{W}_2(\rho_k, \tilde{\rho}_k) = \liminf_{k \rightarrow \infty} \mathbf{W}(u_k, \tilde{u}_k),$$

which was to be proved. Henceforth, one has for all  $t \geq 0$ ,  $h > 0$  and  $k \in \mathbb{N}$  that

$$\mathbf{W}(u_{\tau_k}|_{\mathbb{B}_R(0)}(t+h), u_{\tau_k}|_{\mathbb{B}_R(0)}(t)) \leq \mathbf{W}_2(u_{\tau_k}(t+h), u_{\tau_k}(t)). \tag{4.14}$$

Hypothesis (2.1) from Theorem 2.3 is obviously satisfied thanks to Lemma 3.8(c). Concerning hypothesis (2.2) from Theorem 2.3, we claim that

$$\begin{aligned} & \sup_{k \in \mathbb{N}} \int_0^{T-h} \mathbf{W}_2(u_{\tau_k}(t+h), u_{\tau_k}(t)) \, dt \\ & \leq \max\left(1, \sqrt{T+\bar{\tau}}\right) \sqrt{2C(T+\bar{\tau})(\mathcal{E}(u_0) - \inf \mathcal{E})(T+\bar{\tau})h}, \end{aligned} \quad (4.15)$$

for all  $h \in (0, \bar{\tau})$ , from which (2.2) follows using (4.14). Indeed, for fixed  $k \in \mathbb{N}$  and  $h \in (0, \tau_k]$ , one has, for  $N := \left\lfloor \frac{T}{\tau_k} \right\rfloor$ ,

$$\begin{aligned} & \int_0^{T-h} \mathbf{W}_2(u_{\tau_k}(t+h), u_{\tau_k}(t)) \, dt = \sqrt{C(T+\bar{\tau})} \sum_{n=1}^N h \mathbf{W}_{m(n\tau_k, \cdot)}(u_{\tau_k}^n, u_{\tau_k}^{n+1}) \\ & \leq \sqrt{2C(T+\bar{\tau})(\mathcal{E}(u_0) - \inf \mathcal{E})} \sqrt{h^2 N} \\ & \leq \sqrt{2(\mathcal{E}(u_0) - \inf \mathcal{E})(T+\bar{\tau})h}, \end{aligned}$$

thanks to Hölder's inequality and the total square distance estimate (3.3). On the other hand, for  $h \in (\tau_k, \bar{\tau}]$ , we directly get from the Hölder type estimate (3.4):

$$\begin{aligned} & \int_0^{T-h} \mathbf{W}_2(u_{\tau_k}(t+h), u_{\tau_k}(t)) \, dt \leq (T-h) \sqrt{2C(T+\bar{\tau})(\mathcal{E}(u_0) - \inf \mathcal{E})h} \\ & \leq (T+\bar{\tau}) \sqrt{2C(T+\bar{\tau})(\mathcal{E}(u_0) - \inf \mathcal{E})h}. \end{aligned}$$

Theorem 2.3 thus is applicable for the family  $(u_{\tau_k}|_{\mathbb{B}_R(0)})_{k \in \mathbb{N}}$ ; hence,  $u_{\tau_k}(t)$  converges to  $u(t)$  in  $L^2(\mathbb{B}_R(0))$ , in measure w.r.t.  $t \in [0, T]$ . By the dominated convergence theorem and the uniform estimate in  $L^\infty([0, T]; L^2(\mathbb{R}^d))$  from Lemma 3.8(a),  $u_{\tau_k}$  converges to  $u$  in  $L^2([0, T] \times \mathbb{B}_R(0))$ , and  $u \in L^\infty([0, T]; L^2(\mathbb{R}^d))$ . After a diagonal argument (considering a family of radii  $R \uparrow +\infty$ ), extracting further subsequences, we have that

$$u_{\tau_k} \rightarrow u \quad \text{in } L^2([0, T]; L^2(\mathbb{B}_R(0))) \text{ for all } R > 0 \text{ simultaneously,} \quad (4.16)$$

$$u_{\tau_k}(t) \rightarrow u(t) \quad \text{in } L^2(\mathbb{B}_R(0)) \text{ for all } R > 0 \text{ simultaneously and almost all } t \in [0, T], \quad (4.17)$$

$$u_{\tau_k}(t, x) \rightarrow u(t, x) \quad \text{for almost all } (t, x) \in [0, T] \times \mathbb{R}^d. \quad (4.18)$$

Fix now a  $t \in [0, T]$  such that (4.17) holds. Since by construction of  $u_{\tau_k}$ , one has

$$u_{\tau_k}(t, x) \leq S\left(\left\lceil \frac{t}{\tau_k} \right\rceil \tau_k\right),$$

for almost all  $x \in \mathbb{R}^d$ , and  $S$  is continuous thanks to condition (M2), one also has

$$u(t, x) - S(t) \leq \lim_{k \rightarrow \infty} \left( u(t, x) - u_{\tau_k}(t, x) + S\left(\left\lceil \frac{t}{\tau_k} \right\rceil \tau_k\right) - S(t) \right) = 0,$$

for almost all  $x \in \mathbb{R}^d$ . Hence,  $u(t, x) \in [0, S(t)]$  a.e.

It now remains to prove that the limit map  $u$  is a solution to (1.3) in the sense of distributions, i.e.: For all  $\eta \in C_c^\infty(\mathbb{R}^d)$  and all  $\psi \in C_c^\infty((0, \infty)) \cap C(\mathbb{R}_{\geq 0})$ , it holds that

$$\begin{aligned} & \int_0^\infty \int_{\mathbb{R}^d} [-\partial_t \psi(t) u(t, x) \eta(x) + \psi(t) m(t, u(t, x)) [\nabla u(t, x) f''(u(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x)] \, dx \, dt \\ & = 0, \end{aligned} \quad (4.19)$$

Since  $t \mapsto \frac{\psi_{\tau_k}(t) - \psi_{\tau_k}(t + \tau_k)}{\tau_k}$  converges uniformly in  $\mathbb{R}_{\geq 0}$  to  $-\partial_t \psi$  and  $u_{\tau_k}$  converges weakly to  $u$ , we immediately have that

$$\lim_{k \rightarrow \infty} \int_0^\infty \int_{\mathbb{R}^d} \frac{\psi_{\tau_k}(t) - \psi_{\tau_k}(t + \tau_k)}{\tau_k} u_{\tau_k}(t, x) \eta(x) \, dx \, dt = \int_0^\infty \int_{\mathbb{R}^d} -\partial_t \psi(t) u(t, x) \eta(x) \, dx \, dt.$$

Thus, by the discrete weak formulation (4.2), it remains to show that (possibly extracting another subsequence) one has

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_0^\infty \int_{\mathbb{R}^d} \left[ \psi_{\tau_k}(t) m \left( \left\lceil \frac{t}{\tau_k} \right\rceil \tau_k, u_{\tau_k}(t, x) \right) [\nabla u_{\tau_k}(t, x) f''(u_{\tau_k}(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \right. \\ \left. - \psi(t) m(t, u(t, x)) [\nabla u(t, x) f''(u(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \right] dx dt = 0. \end{aligned} \quad (4.20)$$

First, thanks to the convergence almost everywhere from (4.18) and the continuity of  $m$  and  $f''$ ,

$$\begin{aligned} \psi_{\tau_k}(t) m \left( \left\lceil \frac{t}{\tau_k} \right\rceil \tau_k, u_{\tau_k}(t, x) \right) [f''(u_{\tau_k}(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \\ - \psi(t) m(t, u(t, x)) [f''(u(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \xrightarrow{k \rightarrow \infty} 0, \end{aligned}$$

a.e. on  $[0, T] \times \mathbb{B}_R(0) \supset \text{supp } \psi \times \text{supp } \eta$ . Furthermore, due to the bounds on  $f''$  and  $m$  (recall Proposition 2.4(b)), one has

$$\begin{aligned} \int_0^\infty \int_{\mathbb{R}^d} \left| \psi_{\tau_k}(t) m \left( \left\lceil \frac{t}{\tau_k} \right\rceil \tau_k, u_{\tau_k}(t, x) \right) [f''(u_{\tau_k}(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \right. \\ \left. - \psi(t) m(t, u(t, x)) [f''(u(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \right|^2 \\ \leq C \int_0^\infty \int_{\mathbb{R}^d} \mathbf{1}_{\text{supp } \psi \times \text{supp } \eta}(t, x) (1 + u_{\tau_k}(t, x)^2 + u(t, x)^2) dx dt, \end{aligned}$$

for some  $C > 0$ , and the right-hand side is uniformly integrable since  $u_{\tau_k} \rightarrow u$  in  $L^2([0, T]; L^2(\mathbb{B}_R(0)))$ , cf. (4.16). Thus, Vitali's convergence theorem yields that

$$\begin{aligned} \int_0^\infty \int_{\mathbb{R}^d} \left| \psi_{\tau_k}(t) m \left( \left\lceil \frac{t}{\tau_k} \right\rceil \tau_k, u_{\tau_k}(t, x) \right) [f''(u_{\tau_k}(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \right. \\ \left. - \psi(t) m(t, u(t, x)) [f''(u(t, x)) + \nabla \phi(x)] \cdot \nabla \eta(x) \right|^2 \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Recalling the weak convergence of  $\nabla u_{\tau_k}$  to  $\nabla u$  in  $L^2([0, T] \times \mathbb{R}^d)$ , we consequently infer (4.20) and also the desired distributional formulation (4.19).  $\square$

The proof of Theorem 1.6 goes along similar lines; we now highlight the important differences.

*Proof of Theorem 1.6:* Our strategy for proving strong convergence of  $u_{\tau_k} \rightarrow u$  in the space  $L^2([0, T]; H^1(\mathbb{B}_R(0)))$  is adapted by choosing the space  $\mathbf{Y} := H^1(\mathbb{B}_R(0))$  and the functional  $\mathcal{A} : \mathbf{Y} \rightarrow [0, \infty]$  defined via

$$\mathcal{A}(u) := \begin{cases} \|u\|_{H^2}^2 & \text{if } u \in H^2(\mathbb{B}_R(0)), \\ +\infty & \text{otherwise.} \end{cases}$$

The main arguments stay—mutatis mutandis—the same. To show that  $u$  is a weak solution to (1.5), one needs to verify that

$$\begin{aligned} \int_0^\infty \int_{\mathbb{R}^d} \left[ \psi_{\tau_k} \text{tr} \left\{ \begin{pmatrix} \mathbb{1}_d \text{div } q_{\tau_k} & q_{\tau_k} \end{pmatrix} \nabla_{(p,z)}^2 f(\nabla u_{\tau_k}, u_{\tau_k}) \begin{pmatrix} \nabla^2 u_{\tau_k} \\ (\nabla u_{\tau_k})^T \end{pmatrix} \right\} \right. \\ \left. - \psi \text{tr} \left\{ \begin{pmatrix} \mathbb{1}_d \text{div } q & q \end{pmatrix} \nabla_{(p,z)}^2 f(\nabla u, u) \begin{pmatrix} \nabla^2 u \\ (\nabla u)^T \end{pmatrix} \right\} \right] dx dt \xrightarrow{k \rightarrow \infty} 0, \end{aligned}$$

for  $q := m(t, u) \nabla \eta$ , or that, equivalently, for each  $i \in \{1, \dots, d\}$ :

$$\int_0^\infty \int_{\mathbb{R}^d} \left[ \psi_{\tau_k} \mathbf{e}_i \begin{pmatrix} \mathbb{1}_d \text{div } q_{\tau_k} & q_{\tau_k} \end{pmatrix} \nabla_{(p,z)}^2 f(\nabla u_{\tau_k}, u_{\tau_k}) \begin{pmatrix} \partial_{x_i} \nabla u_{\tau_k} \\ \partial_{x_i} u_{\tau_k} \end{pmatrix} \right]$$

$$- \psi \mathbf{e}_i \left( \mathbb{1}_d \operatorname{div} q \quad q \right) \nabla_{(p,z)}^2 f(\nabla u, u) \left( \frac{\partial_{x_i} \nabla u}{\partial_{x_i} \nabla u} \right) \Big] dx dt \xrightarrow{k \rightarrow \infty} 0,$$

In view of the weak convergence of  $u_{\tau_k} \rightharpoonup u$  in  $L^2([0, T]; H^2(\mathbb{R}^d))$ , it suffices to show that

$$\psi_{\tau_k} \mathbf{e}_i \left( \mathbb{1}_d \operatorname{div} q_{\tau_k} \quad q_{\tau_k} \right) \nabla_{(p,z)}^2 f(\nabla u_{\tau_k}, u_{\tau_k}) \xrightarrow{k \rightarrow \infty} \psi \mathbf{e}_i \left( \mathbb{1}_d \operatorname{div} q \quad q \right) \nabla_{(p,z)}^2 f(\nabla u, u) \quad (4.21)$$

in  $L^2([0, T]; L^2(\mathbb{R}^d; \mathbb{R}^d))$ . Thanks to the continuity of  $f$ ,  $m$  and  $\partial_z m$  and the strong convergence of  $u_{\tau_k} \rightarrow u$  in  $L^2([0, T]; H^2(\mathbb{B}_R(0)))$  for  $\mathbb{B}_R(0) \supset \operatorname{supp} \eta$ , (4.21) holds pointwise almost everywhere on  $[0, T] \times \mathbb{B}_R(0)$ , since

$$\operatorname{div} q_{\tau_k} = \partial_z m \left( \left[ \frac{t}{\tau} \right] \tau_k, u_{\tau_k} \right) \nabla u_{\tau_k} \nabla \eta + m \left( \left[ \frac{t}{\tau} \right] \tau_k, u_{\tau_k} \right) \Delta \eta.$$

By employing the elementary bounds on  $m$  and  $\partial_z m$  from Proposition 2.4(b), we deduce bounds for  $q_{\tau_k}$  analogous to those in (4.10). Consequently, recalling the bound on  $\nabla_{(p,z)}^2 f$  from (E2), we get

$$\begin{aligned} & \int_0^\infty \int_{\mathbb{R}^d} \left| \psi_{\tau_k} \mathbf{e}_i \left( \mathbb{1}_d \operatorname{div} q_{\tau_k} \quad q_{\tau_k} \right) \nabla_{(p,z)}^2 f(\nabla u_{\tau_k}, u_{\tau_k}) \right|^2 dx dt \\ & \leq C \int_0^\infty \int_{\mathbb{R}^d} \mathbf{1}_{\operatorname{supp} \psi \times \operatorname{supp} \eta}(t, x) (1 + u_{\tau_k}(t, x)^2 + |\nabla u_{\tau_k}(t, x)|^2) dx dt, \end{aligned}$$

and the right-hand side is uniformly integrable. Vitali's theorem thus yields the desired convergence (4.21) in  $L^2([0, T]; L^2(\mathbb{R}^d; \mathbb{R}^d))$ .  $\square$

## 5. GENERALIZATIONS

We give some concluding remarks about possible generalizations of our Theorems 1.5 and 1.6.

**5.1. Bounded spatial domains.** In principle, our results also hold if the spatial variable  $x$  is confined to a bounded and convex domain  $\Omega \subset \mathbb{R}^d$  with sufficiently smooth boundary  $\partial\Omega$ , when we consider the initial-boundary-value problem for (1.1) with the no-flux and Neumann boundary conditions

$$m(t, u(t, \bar{x})) \partial_\nu \frac{\delta \mathcal{E}}{\delta u}(u(t, \bar{x})) = 0 = \partial_\nu u(t, \bar{x}), \quad (5.1)$$

for all  $t > 0$  and  $\bar{x} \in \partial\Omega$ , where  $\nu$  denotes the unit normal vector field to  $\partial\Omega$ .

In this framework, one may replace the conditions (E1) and (E2) on the free energy functional by the following different, mostly weaker assumptions.

**Definition 5.1** (Admissible energy functionals on bounded spatial domains).

(E1') *Second order equations* (1.3): Let  $\mathcal{E}$  be of the form

$$\mathcal{E}(u) = \int_\Omega [f(u(x)) + \phi(x)u(x)] dx,$$

where  $\phi \in C^2(\overline{\Omega})$  and  $f \in C^2(\mathbb{R}_{\geq 0})$  with  $f''(z) \in [\gamma_0, \gamma_1]$  for  $0 < \gamma_0 \leq \gamma_1 < \infty$  and all  $z \geq 0$ .

(E2') *Fourth order equations* (1.5): Let  $\mathcal{E}$  be of the form

$$\mathcal{E}(u) = \int_\Omega [f(\nabla u(x), u(x)) + \phi(x)u(x)] dx,$$

with  $\phi$  as in (E1'). For  $f \in C^2(\mathbb{R}^d \times \mathbb{R}_{\geq 0})$ , we assume the following:

There exist  $0 < \gamma_0 \leq \gamma_1 < \infty$  such that for each  $(p, z) \in \mathbb{R}^d \times \mathbb{R}_{\geq 0}$  and all  $(v_p, v_z) \in \mathbb{R}^d \times \mathbb{R}$ , one has

$$\gamma_0 |v_p|^2 \leq \begin{pmatrix} v_p \\ v_z \end{pmatrix}^\top \nabla_{(p,z)}^2 f(p, z) \begin{pmatrix} v_p \\ v_z \end{pmatrix} \leq \gamma_1 (|v_p|^2 + v_z^2),$$

i.e.  $f$  is convex w.r.t.  $(p, z)$  and uniformly convex w.r.t.  $p$ .

Additionally, we require the following *isotropy condition* w.r.t. the gradient variable  $p$ : There exists a map  $\tilde{f} \in C^2(\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0})$  with  $\partial_r \tilde{f}(r, z) \geq 0$  for all  $(r, z) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$  such that  $f(p, z) = \tilde{f}(|p|, z)$  for all  $(p, z) \in \mathbb{R}^d \times \mathbb{R}_{\geq 0}$ .

Notice that one can in the case (E2') now also allow for integrands  $f$  which do not depend on  $z$ . Since  $\Omega$  is bounded, a control on the gradient  $\nabla u$  in  $L^2$  also yields bounds in  $L^2$  on  $u$  itself via Poincaré's inequality. Clearly, we also do not have to require finiteness of second moments  $\mathfrak{m}_2(u)$  for probability measures  $u \in \mathcal{P}(\Omega)$  on  $\Omega$  explicitly anymore. However, due the appearance of boundary terms, we need the additional isotropy condition (compare to [18]).

The respective existence results—which can be obtained by essentially the same methods as Theorems 1.5 and 1.6—read as follows.

**Theorem 5.2** (Existence: second order case (E1')). *Assume that the mobility  $m$  satisfies (M1)–(M4) and the energy functional  $\mathcal{E}$  is of the form (E1') and let an initial datum  $u_0 \in \mathbf{X} \cap L^2(\mathbb{R}^d)$  with  $u_0(x) \in [0, S(0)]$  for almost every  $x \in \Omega$  be given. Then, for each  $\tau > 0$ , the map  $u_\tau$  obtained via (1.10) is well-defined. Furthermore, for each vanishing sequence  $\tau_k \searrow 0$  ( $k \rightarrow \infty$ ), there exists a (nonrelabelled) subsequence and a limit map  $u : \mathbb{R}_{\geq 0} \rightarrow \mathbf{X}$  such that the following is true for each fixed  $T > 0$ :*

- (a)  $u \in L^\infty([0, T]; L^2(\Omega)) \cap L^2([0, T]; H^1(\Omega)) \cap C^{1/2}([0, T]; \mathbf{W}_2)$ , and for almost every  $t \geq 0$ , one has  $u(t, x) \leq S(t)$  for almost all  $x \in \Omega$ ;
- (b)  $u_{\tau_k}(t) \rightarrow u(t)$  uniformly in  $t \in [0, T]$  with respect to the distance  $\mathbf{W}_2$  on  $\mathcal{P}(\Omega)$ ;
- (c)  $u_{\tau_k}$  converges to  $u$  strongly in  $L^2([0, T]; L^2(\Omega))$  and weakly in  $L^2([0, T]; H^1(\Omega))$ ;
- (d)  $u$  is a solution to (1.3) in the sense of distributions, it attains the boundary conditions (5.1) and the initial condition  $u(0, \cdot) = u_0$  holds a.e. on  $\Omega$ .

In the fourth order case, we have:

**Theorem 5.3** (Existence: fourth order case (E2')). *Assume that the mobility  $m$  satisfies (M1)–(M4) and the energy functional  $\mathcal{E}$  is of the form (E2') and let an initial datum  $u_0 \in \mathbf{X} \cap H^1(\Omega)$  with  $u_0(x) \in [0, S(0)]$  for almost every  $x \in \Omega$  be given. Then, for each  $\tau > 0$ , the map  $u_\tau$  obtained via (1.10) is well-defined. Furthermore, for each vanishing sequence  $\tau_k \searrow 0$  ( $k \rightarrow \infty$ ), there exists a (nonrelabelled) subsequence and a limit map  $u : \mathbb{R}_{\geq 0} \rightarrow \mathbf{X}$  such that the following is true for each fixed  $T > 0$ :*

- (a)  $u \in L^\infty([0, T]; H^1(\Omega)) \cap L^2([0, T]; H^2(\Omega)) \cap C^{1/2}([0, T]; \mathbf{W}_2)$ , and for almost every  $t \geq 0$ , one has  $u(t, x) \leq S(t)$  for almost all  $x \in \Omega$ ;
- (b)  $u_{\tau_k}(t) \rightarrow u(t)$  uniformly in  $t \in [0, T]$  with respect to the distance  $\mathbf{W}_2$  on  $\mathcal{P}(\Omega)$ ;
- (c)  $u_{\tau_k}$  converges to  $u$  strongly in  $L^2([0, T]; H^1(\Omega))$  and weakly in  $L^2([0, T]; H^2(\Omega))$ ;
- (d)  $u$  is a solution to (1.3) in the sense of distributions, it attains the boundary conditions (5.1) and the initial condition  $u(0, \cdot) = u_0$  holds a.e. on  $\Omega$ .

**5.2. Non-Lipschitz mobility functions.** A second possibility of extension is concerned with the assumptions (M1)–(M4) on the mobility function  $m$ . Specifically, we consider mobilities which satisfy (M1), (M2) and (M4) but not the Lipschitz-semiconcavity condition (M3). In this framework, we restrict ourselves to the case of a bounded spatial domain  $\Omega \subset \mathbb{R}^d$ , as in Section 5.1. Our strategy is as follows: We approximate  $m$  similarly to [17] with mobilities  $m_\delta$ —which satisfy (M3)—for small  $\delta > 0$  (see Definition 5.6 below). The respective family of weak solutions  $(u_\delta)_{\delta > 0}$  provided by the Theorems 5.2 and 5.3 for  $m_\delta$  in place of  $m$  is then expected to converge to a limit  $u$  which is a weak solution of the original problem. The precise assumptions read as follows:

**Definition 5.4** (Non-Lipschitz mobilities).

- (M1')  $m$  satisfies condition (M1) for a *constant* function  $S(\cdot)$  (with value  $S \in \mathbb{R}_{>0} \cup \{+\infty\}$ ).
- (M2') Condition (M2) holds.
- (M3') There exists a map  $\tilde{m} : [0, S] \rightarrow \mathbb{R}$  satisfying (M1) with the property

$$m(t, z) \leq \tilde{m}(z) \quad \text{for all } t \geq 0 \text{ and all } z \in [0, S].$$

(M4') Condition (M4) holds.

When considering fourth-order equations (cf. (E2')), due to the appearance of gradients, the following condition on  $\partial_z m$  is needed:

(M5') One has for each  $t \geq 0$  that

$$\lim_{z \searrow 0} \partial_z m(t, z) \sqrt{z} = 0,$$

and, if  $S < \infty$ , also

$$\lim_{z \nearrow S} \partial_z m(t, z) \sqrt{S - z} = 0,$$

Notice that we do not require that the auxiliary time-independent function  $\tilde{m}$  satisfies (M3): The distance  $\mathbf{W}_{\tilde{m}}$  acts as a surrogate for the  $L^2$ -Wasserstein distance  $\mathbf{W}_2$  for the Hölder type estimate (3.4) which is needed in the proof of strong convergence. We restricted our considerations on constant value spaces  $[0, S]$  mostly due to technical reasons: for the approximation in Definition 5.6 below, it is convenient to have equal support for all  $m(t, \cdot)$ . Condition (M5') in particular ensures the convergence of  $\nabla m(t, u_\delta)$  to the respective limit; using the well-known *Lions-Villani estimate* on square roots (see [17] for the detailed calculation) provides the necessary a priori estimate.

The following examples are paradigmatic; they correspond to time-dependent versions of the examples discussed in [8, 16, 17].

**Example 5.5** (Paradigmatic choices).

(a) For  $S = +\infty$  and a sufficiently regular map  $\alpha$  with values in  $(\frac{1}{2}, 1]$ , set

$$m(t, z) = z^{\alpha(t)} \quad \text{for all } t \geq 0 \text{ and } z \geq 0.$$

Choosing the *Dirichlet* energy

$$\mathcal{E}_D(u) = \frac{1}{2} \int_{\Omega} |\nabla u(x)|^2 dx, \quad (5.2)$$

which now is admissible in (E2'), (1.5) reads as the thin film equation with time-dependent mobility:

$$\partial_t u = -\operatorname{div}(u^{\alpha(t)} \nabla \Delta u).$$

(b) For  $S < +\infty$  and sufficiently regular maps  $\sigma_1, \sigma_2$  with values in  $(\frac{1}{2}, 1]$ , set

$$m(t, z) = z^{\sigma_1(t)}(S - z^{\sigma_2(t)}) \quad \text{for all } t \geq 0 \text{ and } z \in [0, S].$$

With the conditions (M1')–(M4'), we can—uniformly with respect to time  $t$ —define the approximating mobilities  $m_\delta$  as in [17] distinguishing the cases  $S < \infty$  and  $S = +\infty$ :

**Definition 5.6** (Approximation of  $m$ ). Let  $\delta \in (0, \bar{\delta})$  and  $\bar{\delta} > 0$  sufficiently small.

If  $S < \infty$ , define for each fixed  $t \geq 0$  and  $z \in [0, S]$ :

$$m_\delta(t, z) := m\left(t, \frac{z_{\delta,1}(t) - z_{\delta,2}(t)}{S} z - z_{\delta,1}(t)\right) - \delta, \quad (5.3)$$

where  $z_{\delta,1}(t) < z_{\delta,2}(t)$  are the two solutions of  $m(t, z) = \delta$  for  $z$  at fixed  $t$ .

If  $S = \infty$ , define

$$m_\delta(t, z) := m(t, z + z_\delta(t)) - \delta, \quad (5.4)$$

for all  $t \geq 0$  and  $z \geq 0$ , where  $z_\delta(t)$  is the unique solution of  $m(t, z) = \delta$  for  $z$  at fixed  $t$ .

Thanks to the assumed regularity of  $m$  w.r.t.  $t$  and the fact that  $S(\cdot)$  remains constant, elementary properties of the approximation in Definition 5.6 carry over from the autonomous case [17] to the non-autonomous case. One has (compare to [17]):



**Lemma 5.7** (Properties of the approximation). *Let  $\bar{\delta} > 0$  be sufficiently small. The following statements hold:*

- (a) *For each  $\delta \in (0, \bar{\delta})$ , the map  $m_\delta$  satisfies (M1)–(M4) and the estimate  $0 \leq m_\delta(t, z) \leq m(t, z)$  for all  $t \geq 0$  and  $z \in [0, S]$ .*
- (b) *For each  $\delta \in (0, \bar{\delta})$ , all  $t \geq 0$  and all  $u, \tilde{u} \in \mathbf{X}(0)$ :*

$$\mathbf{W}_{\tilde{m}}(u, \tilde{u}) \leq \mathbf{W}_{m(t, \cdot)}(u, \tilde{u}) \leq \mathbf{W}_{m_\delta(t, \cdot)}(u, \tilde{u}).$$

- (c) *As  $\delta \searrow 0$ ,  $m_\delta \rightarrow m$  locally uniformly in  $\mathbb{R}_{\geq 0} \times [0, S]$  and  $\partial_z m_\delta \rightarrow \partial_z m$  locally uniformly in  $\mathbb{R}_{\geq 0} \times (0, S)$ .*

Using the methods from [17] and [28], one finds thanks to Lemma 5.7 that the map  $h_\delta$  from assumption (M4) admits for all  $T > 0$  the bound

$$|h_\delta(t, z)| \leq C(z^2 + 1) \quad \text{for all } t \in [0, T] \text{ and } z \in [0, S],$$

for a constant  $C > 0$  which is *independent* from  $\delta \in (0, \bar{\delta})$ . In consequence, the following  $\delta$ -uniform *a priori* estimates hold:

**Lemma 5.8** ( $\delta$ -uniform a priori estimates, case (E1')). *Let  $\bar{\delta} > 0$  sufficiently small and let  $T > 0$ . Then, there exists a constant  $C > 0$  such that for all  $\delta \in (0, \bar{\delta})$  and all  $t \in [0, T]$ :*

$$\begin{aligned} \|u_\delta\|_{L^\infty([0, T]; L^2(\mathbb{R}^d))} &\leq C, \\ \|u_\delta\|_{L^2([0, T]; H^1(\mathbb{R}^d))} &\leq C, \\ \mathbf{W}_{\tilde{m}}(u_\delta(t), u_\delta(s)) &\leq C\sqrt{|t - s|} \quad \text{for all } s, t \in [0, T]. \end{aligned}$$

**Lemma 5.9** ( $\delta$ -uniform a priori estimates, case (E2')). *Let  $\bar{\delta} > 0$  sufficiently small and let  $T > 0$ . Then, there exists a constant  $C > 0$  such that for all  $\delta \in (0, \bar{\delta})$  and all  $t \in [0, T]$ :*

$$\begin{aligned} \|u_\delta\|_{L^\infty([0, T]; H^1(\mathbb{R}^d))} &\leq C, \\ \|u_\delta\|_{L^2([0, T]; H^2(\mathbb{R}^d))} &\leq C, \\ \mathbf{W}_{\tilde{m}}(u_\delta(t), u_\delta(s)) &\leq C\sqrt{|t - s|} \quad \text{for all } s, t \in [0, T]. \end{aligned}$$

With similar arguments as in Section 4.2, one passes to the limit  $u_\delta \rightarrow u$  as  $\delta \searrow 0$ . The required strong convergence can again be achieved via Theorem 2.3: the necessary pseudo-distance  $\mathbf{W}$  is chosen to be the (auxiliary) distance  $\mathbf{W}_{\tilde{m}}$ . Note that since we work in bounded spatial domains  $\Omega \subset \mathbb{R}^d$ , an infimization process as in (4.12) is not needed here. As a last step in the proof, one passes to the limit in the weak formulation of (1.3) or (1.5) for  $m_\delta$  and  $u_\delta$  in place of  $m$  and  $u$ , respectively. Again, similar methods as in Section 4.2 apply; convergence of the mobilities and the mobility gradients can be proved by essentially the same strategy as in [17]. In the end, one arrives at:

**Theorem 5.10** (Existence for non-Lipschitz mobilities: second order case (E1')). *Assume that the mobility  $m$  satisfies (M1')–(M4') and that the energy functional  $\mathcal{E}$  is of the form (E1'). Let an initial datum  $u_0 \in \mathbf{X} \cap L^2(\mathbb{R}^d)$  with  $u_0(x) \in [0, S]$  for almost every  $x \in \Omega$  be given. Define  $m_\delta$  for  $\delta \in (0, \bar{\delta})$  and sufficiently small  $\bar{\delta}$  as in Definition 5.6 and let  $u_\delta$  be a weak solution to (1.3) with  $m_\delta$  in place of  $m$  and initial condition  $u_0$  in the sense of Theorem 5.2, for each  $\delta \in (0, \bar{\delta})$ . Then, there exists a vanishing sequence  $\delta_k \rightarrow 0$  and a map  $u : \mathbb{R}_{\geq 0} \rightarrow \mathbf{X}$  such that for the sequence  $(u_{\delta_k})_{k \in \mathbb{N}}$  and the limit  $u$ , one has*

- (a)  *$u \in L^\infty([0, T]; L^2(\Omega)) \cap L^2([0, T]; H^1(\Omega)) \cap C^{1/2}([0, T]; \mathbf{W}_{\tilde{m}})$ , and  $u(t, x) \leq S$  for almost all  $t \geq 0$  and  $x \in \Omega$ ;*
- (b)  *$u_{\delta_k}(t) \rightarrow u(t)$  uniformly in  $t \in [0, T]$  with respect to the distance  $\mathbf{W}_{\tilde{m}}$  on  $\mathbf{X}(0)$ ;*
- (c)  *$u_{\delta_k}$  converges to  $u$  strongly in  $L^2([0, T]; L^2(\Omega))$  and weakly in  $L^2([0, T]; H^1(\Omega))$ ;*
- (d)  *$u$  is a solution to (1.3) in the sense of distributions, it attains the boundary conditions (5.1) and the initial condition  $u(0, \cdot) = u_0$  holds a.e. on  $\Omega$ .*

Again, the fourth-order case is similar:

**Theorem 5.11** (Existence for non-Lipschitz mobilities: fourth order case (E2')). *Assume that the mobility  $m$  satisfies (M1')–(M5') and that the energy functional  $\mathcal{E}$  is of the form (E2'). Let an initial datum  $u_0 \in \mathbf{X} \cap H^1(\mathbb{R}^d)$  with  $u_0(x) \in [0, S]$  for almost every  $x \in \Omega$  be given. Define  $m_\delta$  for  $\delta \in (0, \bar{\delta})$  and sufficiently small  $\bar{\delta}$  as in Definition 5.6 and let  $u_\delta$  be a weak solution to (1.5) with  $m_\delta$  in place of  $m$  and initial condition  $u_0$  in the sense of Theorem 5.3, for each  $\delta \in (0, \bar{\delta})$ . Then, there exists a vanishing sequence  $\delta_k \rightarrow 0$  and a map  $u : \mathbb{R}_{\geq 0} \rightarrow \mathbf{X}$  such that for the sequence  $(u_{\delta_k})_{k \in \mathbb{N}}$  and the limit  $u$ , one has*

- (a)  $u \in L^\infty([0, T]; H^1(\Omega)) \cap L^2([0, T]; H^2(\Omega)) \cap C^{1/2}([0, T]; \mathbf{W}_{\bar{m}})$ , and  $u(t, x) \leq S$  for almost all  $t \geq 0$  and  $x \in \Omega$ ;
- (b)  $u_{\delta_k}(t) \rightarrow u(t)$  uniformly in  $t \in [0, T]$  with respect to the distance  $\mathbf{W}_{\bar{m}}$  on  $\mathbf{X}(0)$ ;
- (c)  $u_{\delta_k}$  converges to  $u$  strongly in  $L^2([0, T]; H^1(\Omega))$  and weakly in  $L^2([0, T]; H^2(\Omega))$ ;
- (d)  $u$  is a solution to (1.5) in the sense of distributions, it attains the boundary conditions (5.1) and the initial condition  $u(0, \cdot) = u_0$  holds a.e. on  $\Omega$ .

#### REFERENCES

- [1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [2] B. Aulbach and N. V. Minh. Nonlinear semigroups and the existence and stability of solutions of semilinear nonautonomous evolution equations. *Abstr. Appl. Anal.*, 1(4):351–380, 1996.
- [3] J.-P. Bartier, J. Dolbeault, R. Illner, and M. Kowalczyk. A qualitative study of linear drift-diffusion equations with time-dependent or degenerate coefficients. *Math. Models Methods Appl. Sci.*, 17(3):327–362, 2007.
- [4] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.
- [5] A. Blanchet, J. A. Carrillo, D. Kinderlehrer, M. Kowalczyk, P. Laurençot, and S. Lisini. A hybrid variational principle for the Keller-Segel system in  $\mathbb{R}^2$ . *ESAIM Math. Model. Numer. Anal.*, 49(6):1553–1576, 2015.
- [6] J. A. Carrillo, S. Lisini, G. Savaré, and D. Slepčev. Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.*, 258(4):1273–1309, 2010.
- [7] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.
- [8] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.
- [9] L. C. F. Ferreira and J. C. Valencia-Guevara. Gradient flows of time-dependent functionals in metric spaces and applications for PDEs, 2015. Preprint. arXiv:1509.04161.
- [10] U. Gianazza, G. Savaré, and G. Toscani. The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. *Arch. Ration. Mech. Anal.*, 194(1):133–220, 2009.
- [11] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [12] P. Laurençot and B.-V. Matioc. A gradient flow approach to a thin film approximation of the Muskat problem. *Calc. Var. Partial Differential Equations*, 47(1-2):319–341, 2013.
- [13] D. Lengeler and T. Müller. Scalar conservation laws on constant and time-dependent Riemannian manifolds. *J. Differential Equations*, 254(4):1705–1727, 2013.
- [14] M. Liero and A. Mielke. Gradient structures and geodesic convexity for reaction-diffusion systems. *Phil. Trans. R. Soc. A*, 371(2005):20120346/1–20120346/28, 2013.
- [15] S. Lisini. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.*, 15(3):712–740, 2009.

- [16] S. Lisini and A. Marigonda. On a class of modified Wasserstein distances induced by concave mobility functions defined on bounded intervals. *Manuscripta Math.*, 133(1-2):197–224, 2010.
- [17] S. Lisini, D. Matthes, and G. Savaré. Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics. *J. Differential Equations*, 253(2):814–850, 2012.
- [18] D. Loibl, D. Matthes, and J. Zinsl. Existence of weak solutions to a class of fourth order partial differential equations with Wasserstein gradient structure, 2015. Preprint. arXiv:1507.05507.
- [19] D. Matthes, R. J. McCann, and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12):1352–1397, 2009.
- [20] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.
- [21] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
- [22] L. Petrelli and A. Tudorascu. Variational principle for general diffusion problems. *Appl. Math. Optim.*, 50(3):229–257, 2004.
- [23] S. Plazotta and J. Zinsl. High-frequency limit of non-autonomous gradient flows, 2016. Preprint. arXiv:1601.04445.
- [24] R. Rossi and G. Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(2):395–431, 2003.
- [25] K.-T. Sturm. Super-Ricci flows for metric measure spaces. I, 2016. Preprint. arXiv:1603.02193.
- [26] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [27] J. Zinsl. Existence of solutions for a nonlinear system of parabolic equations with gradient flow structure. *Monatsh. Math.*, 174(4):653–679, 2014.
- [28] J. Zinsl. The gradient flow of a generalized Fisher information functional with respect to modified Wasserstein distances, 2016. Preprint. arXiv:1603.01375.
- [29] J. Zinsl and D. Matthes. Transport distances and geodesic convexity for systems of degenerate diffusion equations. *Calc. Var. Partial Differential Equations*, 54(4):3397–3438, 2015.

ZENTRUM FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT MÜNCHEN, 85747 GARCHING, GERMANY  
 E-mail address: zinsl@ma.tum.de